NICOLA BRUNO
FRANCESCO PAVANI
MASSIMILIANO ZAMPINI

La percezione multisensoriale

il Mulino

I lettori che desiderano informarsi sui libri e sull'insieme delle attività della Società editrice il Mulino possono consultare il sito Internet:
www.mulino.it

ISBN 88-15-13852-1

Indice

Introduzione

Tradizionalmente, lo studio della percezione coincide con lo studio della percezione visiva, o la percezione acustica, tattile, olfattiva, e così via. In questo libro adottiamo una prospettiva diversa: una prospettiva *multisensoriale*. Il nostro punto di partenza è semplice. Nel mondo in cui viviamo le informazioni su oggetti ed eventi sono molto spesso disponibili a più canali sensoriali. Questa semplice constatazione porta a mettere in discussione il concetto tradizionale di modalità di senso, in favore di un approccio in cui i processi percettivi vengono studiati come sistemi attivi e interconnessi fra loro piuttosto che come sensi passivi e indipendenti.

1. QUANTI SONO I SENSI?

Il mondo in cui viviamo è ricco di proprietà percepibili e queste proprietà sono potenzialmente disponibili a molti **canali sensoriali** diversi. Pensiamo, ad esempio, a come potremmo usare la percezione quando andiamo al bar a bere un caffè. Informazioni visive, ma anche informazioni tattili e muscolari, ci possono aiutare a guidare la mano verso la tazzina da caffè per portarla alla bocca. Informazioni legate a proprietà chimiche della sostanza, rilevabili da naso e lingua ma anche dalla visione, possono contribuire a formare l'esperienza del sapore. Informazioni visive, assieme a informazioni provenienti dall'organo dell'equilibrio, dai muscoli e in qualche misura anche dall'udito possono darci un'idea della direzione dei nostri spostamenti dentro e fuori dal bar. Possiamo ricevere informazioni sulle parole pronunciate dalla persona con cui conversia-

mo al bar grazie al canale uditivo ma anche grazie alla visione dei movimenti labiali. Il mondo è quasi sempre multisensoriale, e la natura potenzialmente multisensoriale del mondo in cui viviamo pone una serie di problemi scientifici importanti. In particolare, fino a che punto la percezione sfrutta le informazioni disponibili in diversi canali sensoriali per costruire una rappresentazione del mondo? E se le sfrutta, il risultato di queste **interazioni multisensoriali** è solo la somma delle informazioni elaborate individualmente dai sensi coinvolti, o fa emergere qualcosa di più, una sorta di «valore aggiunto» della multisensorialità? E a che livello di elaborazione dell'informazione avvengono queste interazioni? E infine, quali sono i principi che regolano queste interazioni? Di queste e di molte altre questioni si occupa questo libro. Per comprendere appieno il senso di queste domande, è opportuno iniziare con una riflessione sul concetto stesso di modalità di senso. Un concetto che diamo per scontato, ma che forse tanto scontato non è. Iniziamo allora con una domanda inusuale: siete proprio sicuri che i sensi siano cinque?

L'attribuzione alla mente di cinque sensi ha radici profonde nella nostra cultura e ci appare un elemento fondante di come pensiamo a noi stessi. La storia

fig. 1. Giuseppe Recco (1634-1695), *Natura morta con dolci, fiori e strumenti musicali (I cinque sensi)*. Olio su tela.

dell'arte, ad esempio, è ricca di celebrazioni degli aspetti edonici dell'esperienza sensibile. Invariabilmente questo tipo di opere prendono la forma di un qualche tipo di allegoria dei «cinque» sensi. Come nella figura 1, dove i sensi sono evocati in relazione a oggetti potenzialmente sensibili o in grado di produrre ben definite conseguenze sensoriali: il liuto e il flauto per i suoni, il cannocchiale per la vista, i fiori per l'olfatto, e così via. Nella moderna società della comunicazione, il controllo dell'esperienza mediante la stimolazione dei «cinque» sensi assume un ruolo via via più importante. Si pensi, ad esempio, alla comunicazione aziendale (nella progettazione di punti vendita, o nei messaggi pubblicitari) in cui la manipolazione dei «cinque» sensi per il controllo dell'esperienza di consumo tende ormai a sostituire l'enfasi tradizionale sulle funzioni e il valore dei prodotti [Schmitt 1999]. O ancora a certa pedagogia applicata [Ceriani e Nigro 2006] in cui alla stimolazione simultanea dei «cinque» sensi – o di un loro sottoinsieme – viene attribuito il ruolo di favorire l'apprendimento e la creatività nell'affrontare i problemi, a scuola come in azienda.

Che i sensi siano cinque è dunque qualcosa che la maggior parte di noi dà per scontato. Ma i sensi sono davvero cinque? Qualche visita alle biblioteche, reali e virtuali, riserva non poche sorprese. Aristotele (e dopo di lui tutti i filosofi che si sono ispirati alla tradizione artistotelica) contava sei sensi, non cinque, perché oltre a vista, udito, tatto, gusto e olfatto riteneva necessario ipotizzare un **senso comune** [Gregoric 2007], su cui torneremo nell'epilogo al libro. Secondo lo spiritoso *Libro dell'ignoranza*, tradotto in italiano un paio d'anni fa [Lloyd e Mitchinson 2006], i sensi sono nove: accanto ai cinque sensi tradizionali, andrebbero citati anche il senso della temperatura, quello dell'equilibrio, quello del dolore e la «consapevolezza del corpo». Se invece dobbiamo credere alla voce «sense» dell'enciclopedia online Wikipedia, i sensi sono undici, perché ai cinque sensi tradizionali ne vanno aggiunti altri sei, e precisamente il senso del dolore, dell'equilibrio, della posizione e del movimento corporeo (propriocezione e cinestesia), il senso della temperatura, il senso del tempo e perfino (ma, viene detto, solo in alcuni soggetti) la sensibilità ai campi magnetici.

Consultando qualche manuale di fisiologia, apprendiamo che il concetto di canale sensoriale si fonda su quello di **recettore**, ossia di cellula nervosa deputata a convertire l'energia di uno stimolo esterno in un segnale nervoso. Questo non è un manuale di fisiologia umana, ma proviamo ad applicare questo criterio. Ci sono quattro tipi di recettori nella retina, tre tipi di cono per la visione diurna, e i bastoncelli per la visione notturna. Uno nella coclea, le cellule ciliate uditive. Due nell'organo di Corti e negli organi vestibolari, le cellule ciliate di tipo I e II del «senso» dell'equilibrio. Poi dobbiamo contarne almeno sette variamente

distribuiti sotto la cute, i corpuscoli di Pacini, di Meissner, di Ruffini, i recettori di Merkel, i recettori per il caldo, quelli per il freddo, e infine le terminazioni nervose libere che sembrano avere un ruolo sia nella sensazione di dolore sia in altre sensazioni ancora legate a stimoli meccanici. Inoltre nei muscoli, tendini e articolazioni troviamo i fusi neuromuscolari, gli organi tendinei del Golgi, i recettori articolari. Nella bocca troviamo le papille gustative, che si distinguono in circumvallate, foliate e fungiformi. La loro distribuzione in parti diverse della lingua produce zone di maggior sensibilità ai cosiddetti quattro gusti primari: dolce, amaro, acido e salato. Inoltre, secondo alcuni ricercatori, esisterebbe un quinto tipo di recettore gustativo, responsabile per il gusto primario che i giapponesi chiamano *umami* (per i dettagli su queste e altre complicazioni legate al gusto dei cibi vi rimandiamo al cap. 6). La codifica degli odori da parte delle cellule ciliate olfattive non è ancora ben compresa, ma sembra plausibile che queste cellule vadano classificate in un certo numero di recettori diversi, specializzati per specifici componenti odoranti. Alcuni autori suggeriscono inoltre che nella cavità nasale sia presente, oltre all'epitelio olfattivo, anche un sistema recettoriale (detto organo vomeronasale) specializzato per la percezione di particolari sostanze chimiche che sarebbero in grado di suscitare reazioni fisiologiche o comportamentali specifiche, ad esempio sessuali (i cosiddetti feromoni). Infine, andrebbero ricordati i diversi tipi di recettori che registrano importanti informazioni sullo stato interno del corpo, come i recettori viscerali e polmonari per lo stiramento, i recettori meccanici del peritoneo, i recettori di flusso nell'uretra, quelli di acidità nell'esofago.

Insomma, se ci mettessimo a contare i diversi tipi di recettori dovremmo concludere che i sensi sono più di una trentina. D'altro canto, gli **stimoli** che raggiungono questi molteplici recettori sono in ultima istanza solo di tre tipi: variazioni di energia elettromagnetica (la luce), variazioni nella stimolazione meccanica (ossia la varietà di pressioni, spostamenti e vibrazioni che sentiamo sulla pelle, e il suono che richiede il propagarsi di una variazione di pressione in un mezzo elastico) o variazioni nella stimolazione chimica. Dei recettori è quindi del tutto corretto dire che sono costruiti per rilevare uno di questi tre tipi di energia. Tenendo conto di questo, il fisiologo potrebbe anche affermare, e sarebbe sensato, che i sensi sono solo tre: ottici, meccanici e chimici. Invece un suo collega, notando che alcuni recettori rilevano informazioni esterne al corpo mentre altri sono specializzati per quelle interne a esso, potrebbe anche ridurre il numero a due: sensi esterni e sensi interni.

Dunque anche solo limitandosi a riflettere sui sensi dell'uomo, quanti siano i sensi non è per nulla chiaro. Paradossalmente, l'unica cosa certa che sembra

emergere dalla nostra lista di possibili risposte è che in base a nessun criterio i sensi sono cinque (ma su questo torneremo nel par. 3). Ma ancora non è finita. Fra i commenti a un blog ci è capitato di leggere: «I sensi sono tatto, vista, olfatto, udito, gusto, il sesto senso, il senso del gol, il senso vietato, il senso dell'ironia, e poi ci sono tutti gli ultrasensi degli insetti che noi manco riusciamo a immaginarci».

In questo libro non ci occupiamo di menti non umane (per chi fosse interessato, rimandiamo al saggio di Vallortigara [2000]), ma l'ignoto umorista non ha tutti i torti. Lo studio della percezione in animali diversi da noi ci riserverebbe altre sorprese ancora, come la sensibilità ai campi elettrici degli squali e dell'ornitorinco, quella per la direzione di polarizzazione della luce delle api, o ancora l'ecolocazione dei pipistrelli e dei cetacei.

E allora chi ha ragione? Come molti lettori avranno capito, lo scopo di questa breve discussione non è tanto quello di contare i sensi, ammesso che sia possibile, ma semmai quello di problematizzare una nozione che diamo per scontata, ma che scontata non è. In questo una domanda apparentemente innocente, come quella che ci siamo posti qui, si rivela densa di implicazioni epistemologiche. Innanzi tutto, le diverse maniere in cui è possibile contare i sensi ci fanno apprezzare come, nello studio della mente, sia spesso possibile aggredire i problemi a diversi livelli di analisi. A volerla approfondire, la questione ci porterebbe ad affrontare il problema stesso di cosa costituisca una spiegazione causale in biologia e nelle neuroscienze in particolare [Hogan 1994]. Per i fini del nostro discorso ci basterà distinguere fra il livello del **sostrato neurale** della mente e quello delle sue **funzioni adattive**. Abbiamo adottato il primo dei due livelli di analisi quando abbiamo parlato di sensi in termini di cellule recettoriali o di organi come l'occhio, la coclea, il vestibolo. Abbiamo invece adottato il secondo quando abbiamo ipotizzato di analizzare i sensi in riferimento ai mezzi informativi che l'ambiente rende disponibili al nostro organismo, o ancora quando parlando di senso dell'equilibrio o della posizione ci siamo posti il problema della funzione di questi meccanismi sensoriali nel regolare il comportamento.

Per molti anni la posizione della *psicologia cognitivista* è stata fortemente sbilanciata verso il livello delle funzioni, e in questo sposava quell'approccio alla filosofia della mente che viene appunto chiamato *funzionalismo* [Fodor 1968; Putnam 1975]. Il funzionalismo ritiene che i fenomeni mentali debbano essere compresi studiando il ruolo che essi hanno nel sistema organismo-ambiente, ossia studiando *come* e *per quale scopo* i processi mentali elaborano informazioni e producono comportamenti. Per i funzionalisti, lo studio delle funzioni è indipendente da *che cosa* implementa la funzione, ossia del suo sostrato fisico o biologico: uno stato mentale è quello che è in virtù di un insieme di procedure e

relazioni logiche astratte (come in un programma di computer), non delle strutture concrete che le implementano (siano esse una rete di cellule in un cervello, o un chip di silicio, o qualsiasi altra cosa).

Nelle neuroscienze cognitive contemporanee, a cui lo studio psicologico dei processi cognitivi sempre di più contribuisce, l'assunzione di indipendenza delle funzioni dalle strutture è oggi spesso sostituita da un approccio che enfatizza le relazioni fra questi due livelli di analisi, cercando di armonizzarli (per una discussione del ruolo di diversi livelli di spiegazione nello studio della percezione vedi Marr [1982]). In questa prospettiva, *che cosa* realizza concretamente una funzione pone dei vincoli anche a *come* questa può essere realizzata. Di conseguenza, per spiegare un processo mentale occorre comprendere quale funzione svolge, ma anche a quale sostrato neurale corrisponde. Un ruolo fondamentale in questo senso viene svolto dalle moderne tecniche di *neuroimmagine funzionale*, oltre che al più tradizionale studio dei deficit comportamentali esibiti da pazienti con malattie neurologiche, la *neuropsicologia*. Diamo maggiori dettagli su questi e altri metodi per studiare la mente multisensoriale nel paragrafo 4. Inoltre, un ruolo importante nell'armonizzare i due livelli di spiegazione viene svolto dall'analisi dello *scopo* di un processo mentale nel contesto della storia evolutiva di un organismo. Come dire: i processi mentali sono un insieme di *specializzazioni adattive*, che servono a consentirci di funzionare con successo nel nostro ambiente (altrimenti, la nostra specie si sarebbe estinta) cercando di utilizzare al meglio il materiale biologico di cui siamo dotati.

Da un punto di vista evoluzionistico, il vantaggio di avere una mente multisensoriale è ovvio. Se ci sono incidenti o malattie, un organismo può continuare a interagire in modo abbastanza autonomo col mondo anche se diventa cieco, o sordo, o perde la sensibilità tattile sulla superficie del corpo come i pazienti deafferentati che discuteremo nella prima parte del libro. Se invece tutto va bene, utilizzare simultaneamente diverse fonti di informazione sugli oggetti, la loro disposizione spaziale, la loro relazione con il mio corpo rappresenta comunque una maniera di avere «dati» percettivi potenzialmente più robusti e affidabili. Tutto ciò suggerisce dunque la possibilità di studiare la percezione con una prospettiva diversa da quella dei manuali tradizionali, una *prospettiva multisensoriale* in cui i sensi vengono considerati come *sistemi* interconnessi al servizio del comportamento di un organismo che percepisce e agisce in una specifica nicchia ecologica. L'adozione di una prospettiva multisensoriale, i cui principi generali sono presentati nel paragrafo seguente, richiede di mettere in discussione un assunto forte della psicologia cognitivista classica, l'assunzione di *modularità* dei sistemi di input [Fodor 1983], come vedremo fra breve. Le implicazioni di un approccio sistemico e multisensoriale alla percezione sono sviluppate nel paragrafo 3, dove viene presentato il concetto di *sistema percettivo* [Gibson 1966].

2. DALLA MODULARITÀ DEI SENSI ALLA PROSPETTIVA MULTISENSORIALE

Ma se la nozione stessa dei «cinque» sensi è problematica, come abbiamo cercato di mostrare nel paragrafo precedente, allora perché mai la diamo per scontata con così tanta disinvoltura? Inoltre, se la multisensorialità è così pervasiva come abbiamo suggerito, perché ci viene comodo trattare i sensi come distinti gli uni dagli altri tanto nelle nostre conversazioni quotidiane quanto nelle discussioni accademiche? E ancora, perché un qualsiasi testo tradizionale sulla percezione comprende capitoli sulla visione, sull'udito, sul tatto, e così via, ma quasi mai un capitolo specifico sulla multisensorialità? (Fa eccezione il recente manuale di Mather [2009], che comunque riserva al capitolo sulla multisensorialità solo dieci delle sue oltre quattrocento pagine.) Proprio per comprendere i vantaggi di un approccio multisensoriale allo studio della percezione è utile chiedersi innanzi tutto perché tradizionalmente i sensi vengano studiati separatamente, come sistemi di input indipendenti.

Una prima giustificazione sta nel fatto che, nelle prime fasi dell'elaborazione delle informazioni, gli stimoli fisici vengono trasformati in impulsi neurali (la cosiddetta **trasduzione**) da recettori posti in epiteli sensoriali spazialmente distinti (per esempio, la retina per la vista, la coclea per l'udito, la pelle per il tatto). Se l'interesse di un ricercatore riguarda proprio queste operazioni di codifica iniziale dell'informazione, è del tutto sensato studiare questi sistemi recettoriali ognuno per conto suo. Inoltre, le informazioni sensoriali trasdotte dai diversi recettori raggiungono strutture cerebrali specifiche, ad esempio la corteccia visiva primaria (area 17 di Brodmann) o la corteccia somatosensoriale primaria (aree 1, 2 e 3 di Brodmann). Questo suggerisce l'esistenza di aree cerebrali specializzate per l'elaborazione di informazioni provenienti da uno specifico canale sensoriale, le cosiddette **aree sensoriali primarie** (mostrate in fig. 2 insieme alle aree sensoriali secondarie e alle aree motorie).

Una terza giustificazione, infine, sta nel fatto che le regioni corticali apparentemente deputate all'analisi dei segnali provenienti dai recettori esibiscono un'architettura caratteristica e specifica per una modalità di senso. In particolare, grazie al lavoro del famoso fisiologo canadese Wilder Penfield, sappiamo che gli stimoli tattili vengono mappati in una specifica maniera nella corteccia

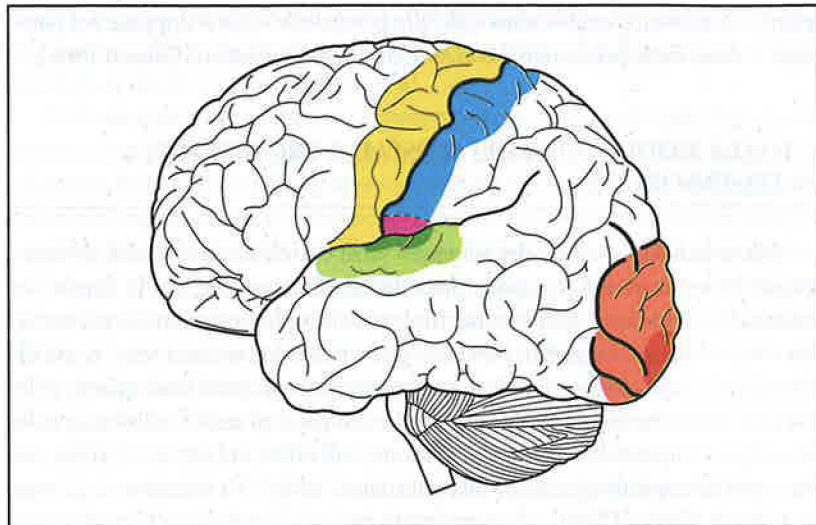



fig. 2. Una rappresentazione schematica delle aree sensoriali primarie e secondarie: rosso, visione; blu, sensazione somatica; fucsia, gusto (solo parzialmente visibile perché si estende in parte anche all'interno della scissura silviana); verde, udito. L'area olfattiva non viene mostrata perché collocata nella zona mediale interna). In giallo l'area motoria primaria.

somatosensoriale, dando luogo a una rappresentazione della superficie corporea che preserva, pur con caratteristiche distorsioni, le relazioni spaziali originali. Infatti questa rappresentazione viene spesso riassunta con il disegno di un corpo umano distorto, un *homunculus somatosensoriale* in cui la proporzione di corteccia dedicata a specifiche parti del corpo corrisponde alla grandezza relativa di quella parte nel disegno stesso. Tralasciando per ora il problema delle distorsioni riscontrabili in queste mappe sensoriali, su cui torneremo nella prima parte del libro, quello che vogliamo sottolineare qui è che questo tipo di rappresentazione utilizza una codifica dello spazio basata su **coordinate somatotopiche** (cioè, su uno schema di riferimento basato sul corpo). Un discorso analogo può essere fatto per la mappatura degli stimoli visivi dalla retina alla corteccia visiva primaria (qui si parla di **coordinate retinotopiche**), o di quelli acustici in cui la mappatura è basata sulla loro frequenza (e infatti si parla di **coordinate tonotopiche**). Anche questo fatto (vedi fig. 3) dunque giustifica l'approccio tradizionale alla percezione, perché suggerisce una forte specializzazione delle regioni corticali deputate all'analisi dei diversi sensi. Scoperte più recenti però sembrano suggerire che le cose siano più complicate di così. Ad esempio, oggi sappiamo che è possibile osservare una convergenza dei diversi canali sensoriali già nelle prime stazioni

sinaptiche della via che porta l'informazione sensoriale dai recettori periferici alla corteccia [Ghazanfar e Schroeder 2006]. Questo suggerisce che l'integrazione multisensoriale potrebbe essere una caratteristica molto più fondamentale di quanto non si ritenesse in passato [Driver e Spence 2000].

Quali che siano i motivi, molti dei ricercatori che si occupano di comprendere la percezione tendono a studiare le singole modalità sensoriali isolando le une dalle altre. La concezione della mente sottostante a queste ricerche è quella della mente modulare [Fodor 1983]. In un libro che ha avuto grande influenza sulla psicologia e sulle neuroscienze cognitive, il filosofo americano Jerry Fodor propose una fondamentale suddivisione tra **facoltà orizzontali**, come la memoria, l'attenzione, il ragionamento, e **facoltà verticali**, come i sistemi di input, (ossia appunto i sensi) e quelli di output (ossia i sistemi motori). Secondo Fodor, il ruolo delle facoltà verticali è trasformare le informazioni percepite in rappresentazioni, che vengono poi fornite alla parte centrale del sistema cognitivo. Fodor riteneva inoltre che la caratteristica fondamentale delle facoltà verticali fosse appunto la modularità. Egli concepiva le facoltà verticali come costituite da sistemi indipendenti con caratteristiche ben precise:

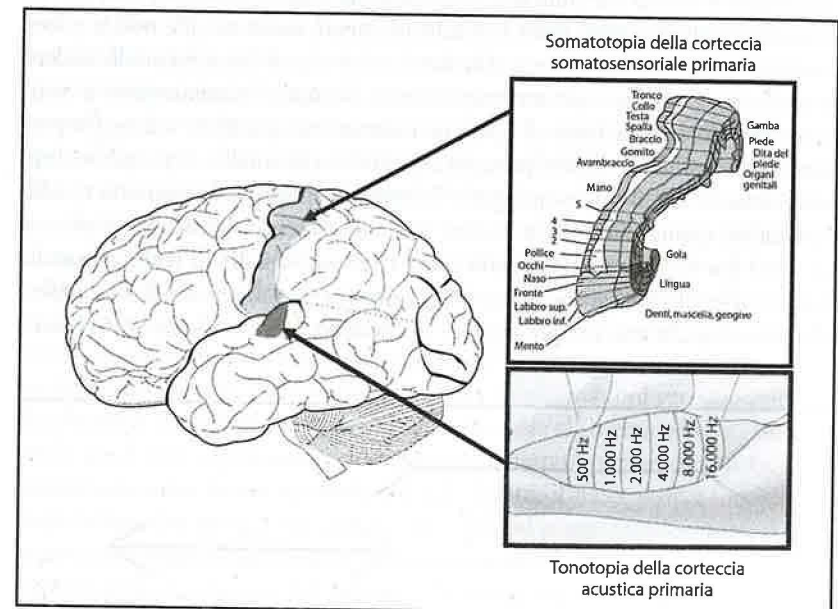



fig. 3. Le mappe somatotopiche e tonotopiche nella corteccia somatosensoriale primaria e nella corteccia uditiva primaria.

1. **specificità per un dominio**: si tratta di strutture altamente specializzate, che possono analizzare solo tipologie di stimoli particolari e specifici per quel modulo;

2. **funzionamento obbligato**: quando sono in presenza del tipo specifico di input che sono deputati ad analizzare, non possono fare a meno di entrare in azione;

3. **accesso centrale limitato**: livelli intermedi di analisi dell'input sono per lo più inaccessibili ai sistemi di controllo centrali;

4. **velocità di funzionamento**;

5. **incapsulamento dell'informazione**: durante il loro funzionamento non hanno accesso né alle rappresentazioni delle conoscenze dell'individuo, né a informazioni comunque provenienti da altre parti del sistema cognitivo dell'individuo.

La teoria di Fodor ha avuto molto successo, anche a causa del fatto che molti meccanismi percettivi sembrano corrispondere bene ai criteri elencati sopra. Pensate alle cosiddette illusioni ottico-geometriche, come l'**illusione di Müller-Lyer** (due segmenti di lunghezza uguale che appaiono diversi quando ci aggiungete le frecce che puntano verso l'interno e verso l'esterno, vedi fig. 4). Questo tipo di situazioni sono straordinariamente resistenti alle nostre conoscenze e aspettative: per quanto ci si sforzi, «non si può fare a meno di» vedere immediatamente i due segmenti come diversi. Dunque, il sistema visivo si comporta proprio come se fosse all'opera un meccanismo specifico, veloce e impermeabile alle influenze di altri processi cognitivi «orizzontali». Si potrebbe dire, naturalmente, che il punto qui non è l'eventuale effetto della memoria o delle aspettative, quanto l'influenza di altre modalità di senso. Di sicuro, se potessi toccare i due segmenti, mi renderei conto che sono uguali? In realtà per molti anni si è ritenuto invece che questa correzione fra segnali sensoriali discordanti non avvenisse. In una celebre ricerca, ad esempio, Rock e Victor [1964] facevano indossare ai soggetti degli occhiali prismatici che restringevano l'immagine di un quadrato sulla retina dei soggetti. Contemporaneamente, i soggetti toccavano il quadrato con una mano. La grandezza percepita però rimaneva sostanzialmente quella prevista in base all'immagine retinica. Insomma, in base a queste ricerche sembrava che la visione della grandezza fosse impermeabile anche alle influenze di un'altra modalità di senso. Proprio come previsto dalla teoria di Fodor.

fig. 4. L'illusione di Müller-Lyer. I due segmenti orizzontali hanno la stessa lunghezza fisica (per un approfondimento del tema delle cosiddette illusioni ottico-geometriche vedi Da Pos e Zambianchi [1996]).

Più di recente però le basi teoriche ed empiriche di questa impostazione sono state fortemente messe in dubbio. In particolare, come già accennato, sembra vacillare l'assunzione che nel cervello vi siano aree sensoriali primarie, specializzate per modalità sensoriale e deputate unicamente all'elaborazione di informazioni provenienti da uno dei «sensi». Studi di neuroanatomia e neurofisiologia su primati non umani e di neuroimmagine nell'uomo hanno permesso di mostrare come l'integrazione multisensoriale avvenga anche in aree che fino a qualche tempo fa venivano considerate unicamente unisensoriali. Per esempio, il gruppo di Gemma Calvert a Oxford ha mostrato che la semplice visione dei movimenti delle labbra di una persona che sta parlando può attivare la corteccia uditiva primaria anche in assenza del suono delle parole prodotte [Calvert *et al.* 1997]. Emiliano Macaluso, Chris Frith e Jon Driver dell'University College di Londra hanno dimostrato che la risposta nella corteccia visiva primaria a stimoli visivi presentati in prossimità di una mano può essere amplificata se le persone contemporaneamente ricevono una stimolazione tattile sulle dita della stessa mano [Macaluso, Frith e Driver 2000]. Già questi risultati sarebbero sufficienti per mettere in crisi un'impostazione strettamente modularista.

È vero che esistono anche esperienze percettive genuinamente unisensoriali, cioè mediate da un unico canale sensoriale. La percezione del colore è solo visiva. Si reagisce al solletico solo se si viene toccati. In molti altri casi però quelle che crediamo esperienze esclusivamente unisensoriali richiedono, in realtà, l'attivazione di un processo multisensoriale. Per esempio, quando si parla con un'altra persona si ha l'impressione che tutte le informazioni vengano veicolate attraverso la voce, e interpretate unicamente grazie al sistema uditivo. Tuttavia, molti studi sulla percezione del linguaggio parlato mostrano che anche le informazioni visive hanno un ruolo centrale. Infatti, il poter vedere i movimenti delle labbra, ma anche l'espressione del volto del nostro interlocutore, ci aiuta a percepire meglio cosa il nostro interlocutore sta dicendo, e in ultima analisi a capire meglio i contenuti del discorso che sta facendo. Persino nell'esempio di percezione visuotattile della grandezza, citato nel paragrafo precedente, risultati recenti hanno dimostrato che la dominanza visiva osservata da Rock e Victor

è solo un caso particolare. Di norma, quando percepiamo la grandezza di un oggetto che vediamo e tocchiamo, il **percetto** è un compromesso tra le informazioni fornite dai due canali [Ernst e Banks 2002].

In tutti i casi come questo, e sono molti, adottare una **prospettiva unisensoriale** ci porta a trascurare alcuni fenomeni percettivi veramente sorprendenti. Prendiamo in considerazione un altro esempio: provate a strofinarvi le mani. La pelle era secca o grassa? Liscia o rugosa? Quale che sia la risposta che avete dato, la sensazione provata sembra genuinamente e unicamente tattile – è localizzata sulla pelle, sembra di sentirla attraverso la pelle. Ma un famoso esperimento ha mostrato che le cose non stanno così [Jousmäki e Hari 1998]. In questo esperimento veniva chiesto ai soggetti di strofinarsi le mani e di dare un giudizio, appunto, su quanto erano lisce e secche. Il suono prodotto dallo strofinio era raccolto da un microfono posto accanto alle mani e veniva fatto sentire alle persone tramite delle cuffie. Il suono presentato poteva essere uguale all'originale, oppure poteva venire modificato, amplificando o attenuando le frequenze più alte (superiori ai 2 Hz). I risultati sono stati davvero sorprendenti. Quando le alte frequenze erano amplificate (percezione acustica) i soggetti riferivano che la pelle sembrava più liscia e asciutta, dava una sensazione simile a quella che dà la pergamena o la carta da forno che si usa per cucinare (e infatti l'effetto viene chiamato *parchment-skin illusion* che potrebbe essere tradotto come «illusione della pelle di pergamena»). Quando invece le alte frequenze erano attenuate i soggetti dicevano di sentire la pelle più ruvida e umida (per conclusioni leggermente diverse, ma comunque coerenti con una interazione audiotattile [Guest e colleghi 2002]). Insomma, come ci sentiamo la pelle è in parte anche una questione di orecchie. Una conclusione che dovrebbe interessare i gestori di profumerie e negozi di articoli cosmetici, e ha già interessato i produttori di alimenti (su questo torneremo nel cap. 6).

E non accade solo per l'udito e il tatto. Ad esempio, quando siete andati al cinema per l'ultima volta? Da dove veniva il suono della voce degli attori sullo schermo? Ovviamente, il suono sembrava venire dalle labbra stesse degli attori. Se siete d'accordo (e presumiamo che lo siate), avete provato un effetto che viene chiamato **illusione del ventriloquo**, la tendenza a localizzare erroneamente la fonte del suono (cioè la voce dell'attore) verso la sua fonte visiva apparente (cioè i movimenti delle labbra sullo schermo) se sussistono certe condizioni che favoriscono questo esito nel processo di integrazione fra i due canali sensoriali [Alais e Burr 2004]. In tutti i casi come questo, dunque, il sistema percettivo integra automaticamente, senza alcun controllo o sforzo da parte nostra, informazioni da più canali. La percezione è ricca di situazioni come queste, e per poterle capire è necessario adottare una **prospettiva multisensoriale**. Capire come funziona la percezione multisensoriale ha importanti conseguenze, sia teoriche sia empiriche, per lo studio di processi cognitivi importanti ma ancora poco studiati o mal compresi, come la percezione del proprio corpo, la psicologia dell'azione intenzionale, la percezione dei cibi e delle bevande di cui ci nutriamo. E il corpus di conoscenze che emerge da queste ricerche ha anche applicazioni rilevanti per la vita quotidiana. Può, per esempio, aiutare ricercatori e designer a progettare prodotti che stimolino più efficacemente tutti i nostri sensi, e che in questo modo contribuiscano a migliorare la qualità di vita. O ancora, in ambito clinico, può contribuire ad alleviare i disturbi di cui soffrono gli individui che hanno subito un'amputazione.

3. DAI SENSI AI SISTEMI PERCETTIVI

Tuttavia adottare una prospettiva multisensoriale non significa solamente constatare che i sensi collaborano fra loro. Nella concezione classica, i sensi funzionano sostanzialmente come *sensori biologici*, cioè come meccanismi in grado di misurare energia fisica di un qualche tipo. Come abbiamo visto, questo livello di analisi è utile per studiare la fisiologia dei recettori, ma mostra dei limiti per capire la percezione come processo cognitivo in senso globale. Il teorico della percezione che ha fornito le basi per il superamento di questa nozione classica dei sensi è stato il grande psicologo americano James J. Gibson. In un libro destinato a diventare uno dei classici della psicologia sperimentale, *The senses considered as perceptual systems* (I sensi considerati come sistemi percettivi), Gibson propose una radicale reinterpretazione del processo percettivo. La tesi del libro, di cui purtroppo non esiste ancora una traduzione in italiano, è ben riassunta da questa citazione: «Dovremo concettualizzare i sensi esterni in una maniera nuova, come attivi piuttosto che passivi, come sistemi piuttosto che come canali, e come interrelati invece che come mutualmente esclusivi» [Gibson 1966, 47]. Proprio per rimarcare il suo nuovo approccio, Gibson propose di abbandonare la terminologia tradizionale e di adottarne una completamente nuova: al livello di analisi della psicologia e delle neuroscienze cognitive, non dobbiamo occuparci di sensi ma di **sistemi percettivi**.

Dove sta la differenza? Come afferma Gibson, in tre cose: la percezione è un processo *attivo*, che va studiato a un livello *sistemico* e con una prospettiva *multisensoriale*. Nel paragrafo precedente abbiamo già discusso in dettaglio l'ultimo dei tre aspetti della proposta gibsoniana: un gran numero di risultati

empirici confermano l'importanza di studiare sistemi percettivi interrelati piuttosto che mutualmente esclusivi. In questo paragrafo completiamo il quadro teorico su cui è basato questo libro con una discussione del primo e del secondo aspetto. Forse il più grande contributo di Gibson è stato notare che lo studio della percezione non può prescindere dalla dimensione temporale. In condizioni naturali, percepire significa esplorare per ottenere informazioni nel corso di attività motorie. Come dice benissimo Gibson in un altro celebre passo, questa constatazione ha una conseguenza fondamentale:

> Un animale vivente non viene stimolato solo dall'ambiente, ma anche da se stesso. Producono stimoli i suoi organi interni, ma anche il movimento delle sue estremità coi loro organi di senso, così come anche la locomozione del suo intero corpo nello spazio. [...] La stimolazione prodotta da un'azione è *ottenuta*, non *imposta*. [...] Essa è intrinseca al fluire dell'attività, non estrinseca a essa; dipende da quel fluire. [...] L'input non è solo afferente, nella terminologia usata dai neurologi, ma riafferente, ossia contingente all'output efferente [*ibidem*, 31; corsivo dell'autore].

Il fatto che le nostre menti siano menti di animali mobili può sembrare ovvio. Ma negli anni '60 e '70 la ricerca disponeva ancora di pochi mezzi per studiare i rapporti fra percezione e azione. Prevalevano le ricerche di tipo **tachistoscopico** (dalle due parole greche che significano «veloce» e «visione»), in cui gli stimoli venivano somministrati per poche frazioni di secondo usando appositi strumenti, detti appunto tachistoscopi, o monitor di computer in grado di controllare con precisione la durata delle schermate. Gibson fu tra i primi a sostenere che questo tipo di ricerche, per quanto interessanti per problemi molto specifici, non erano in grado di dirci molto sulla percezione in condizioni naturali. Si trattava del primo abbozzo di quello che diventerà poi l'**approccio ecologico** alla percezione, delineato da Gibson nel libro che scrisse dieci anni dopo, poco prima di morire [1979; per una discussione aggiornata, vedi Dell'Anna 2006].

In secondo luogo, in quel periodo la ricerca in percezione era ancora dominata da uno schema teorico che affonda le sue radici nella psicofisica fechneriana della seconda metà dell'Ottocento [Fechner 1860], la cosiddetta **catena psicofisica**. Si tratta dell'idea secondo la quale il processo percettivo è riassunto da due passaggi di informazione in sequenza: il primo dallo **stimolo distale** (l'oggetto fisico esterno, indipendente da un atto osservativo) allo **stimolo prossimale** (i recettori nell'organo di senso), il secondo dallo stimolo prossimale al **percetto**

(l'esperienza fenomenica del soggetto). L'elemento cruciale di questo modo di caratterizzare la percezione è che l'informazione viaggia in un solo senso, dal mondo fisico ai recettori grazie a un mezzo informativo (ad esempio, nel caso della visione, la luce riflessa dagli oggetti e rilevata dai recettori), e dai recettori ai sistemi di elaborazione centrale grazie all'elaborazione neurale del segnale afferente. Questo percorso unidirezionale culmina e si conclude poi con la produzione di un percetto cosciente. Gibson è stato uno dei primi a far notare che la catena psicofisica è in realtà solo un pezzo di un processo più articolato, che non è unidirezionale e passivo ma ciclico e attivo. L'informazione afferente proveniente dai recettori viene usata per guidare l'attività esplorativa, grazie alla quale l'organismo rende disponibile nuova informazione che è appunto ri-afferente perché contingente ai segnali efferenti che vengono prodotti per guidare l'esplorazione.

L'idea di un ciclo percezione-azione si è gradualmente affermato nella psicologia cognitiva moderna, grazie anche alla sua adozione da parte di uno dei teorici che erano stati tra i fondatori del cognitivismo prima maniera, Ulric Neisser, in un libro che apparve in Italia addirittura prima del classico di Gibson [Neisser 1976]. La grande maggioranza dei fenomeni multisensoriali che esamineremo in questo libro coinvolgono, oltre alla presenza di informazioni da più canali sensoriali, anche un osservatore attivo e informazioni che variano nel tempo. Questo è vero ad esempio della percezione di come è fatto il nostro corpo, di come viene rappresentato lo spazio, della percezione del sapore dei cibi, e di molte altre situazioni della vita quotidiana. In secondo luogo, alcuni dei fenomeni multisensoriali che esamineremo riguardano direttamente il ruolo della multisensorialità per guidare le azioni. Ad esempio, usiamo continuamente informazioni da canali sensoriali multipli per controllare la postura e dirigere la locomozione. Inoltre le relazioni fra processi motori, processi percettivi e coscienza sono alla base di alcune ricerche recentissime che hanno cercato di gettare luce sulle basi neurali della percezione di sé come **agente intenzionale**. Quello del **ciclo percezione-azione** rappresenta dunque uno schema concettuale indispensabile per capire cosa fa la mente multisensoriale in quasi tutte le situazioni che esamineremo nel libro. Infine, l'adozione dello schema concettuale basato sul ciclo percezione-azione ci suggerisce una soluzione plausibile al paradosso enunciato nel capitolo 1: tutti pensiamo che i sensi siano cinque, ma quando si prova a contarli in base a diversi criteri, tutti apparentemente ragionevoli, i conti non tornano. Il grande fisiologo Henri Piéron [1945] sosteneva che noi pensiamo di avere cinque sensi proprio perché cinque sono le modalità con cui possiamo esplorare attivamente il mondo esterno: guardare, toccare, ascoltare, annusare e assaggiare.

La proposta di Piéron comporta un completo ribaltamento di prospettiva, che non era sfuggito a Gibson. I processi percettivi non sono definiti in funzione dei recettori sensoriali coinvolti, ma in funzione dell'attività esplorativa che siamo in grado di svolgere. Una volta digerito questo passaggio, è naturale abbandonare lo schema della catena psicofisica e adottare quello del ciclo percezione-azione. Una seconda conseguenza è che, come dice sempre Gibson, l'attività percettiva non può essere definita nemmeno in riferimento agli organi che ospitano i recettori: «Evidentemente i veri organi della sensibilità non sono parti del corpo, e le unità dell'anatomia non sono le unità della funzione. Gli organi fanno parte di sistemi» [Gibson 1966, 42]. La parola «sistema» naturalmente ha molti significati. Un **sistema** è, nel linguaggio comune, una maniera di pianificare o strutturare degli oggetti o eventi. A un livello più tecnico, si parla di sistema ogni qual volta un insieme di elementi mostra un qualche tipo di organizzazione (il sistema solare, ad esempio). Nel linguaggio scientifico, infine, la teoria dei sistemi è un approccio multidisciplinare a un gran numero di fenomeni fisici e biologici, fondato dal biologo austriaco Ludwig von Bertalanffy [1968]. Un sistema è, in quest'ultima accezione, un modello di un certo dominio del reale il cui comportamento non è riducibile a quello delle parti che lo compongono, ma è invece caratterizzato dalla sua struttura e dai processi che mettono in relazione fra loro le componenti stesse, secondo specifiche gerarchie. Inoltre, poiché le loro proprietà dipendono dalle relazioni fra le parti, i sistemi possono esibire **proprietà emergenti**, ossia proprietà che sono date dalle relazioni fra le parti e non sono riducibili alle parti stesse prese da sole.

Quest'ultimo significato è quello che aveva in mente Gibson nel proporre la sua nozione di sistemi percettivi, che per lui erano classificabili in base a due criteri: la modalità di acquisizione attiva di informazione, e la classe di oggetti ed eventi nell'ambiente che sono specificati dall'informazione raccolta (vedi tab. 1). Usando questi criteri, Gibson proponeva di distinguere fra un sistema per l'**orientamento di base**, in grado di acquisire automaticamente informazioni per il controllo della postura e sulla direzione della gravità; un sistema **uditivo** per la rilevazione della natura e posizione di oggetti ed eventi esterni; un sistema **aptico** per la rilevazione dei contatti con le superfici e per ottenere informazioni sulla forma e consistenza degli oggetti; un sistema **gustativo/olfattivo** per rilevare il valore nutritivo ed edonico delle sostanze; e infine un sistema **visivo**, per tutte le informazioni veicolabili dalla struttura spaziotemporale della luce, e in particolare per la forma, posizione nello spazio e composizione materiale di oggetti animati e inanimati. Questa classificazione ci appare oggi in qualche misura arbitraria, o comunque troppo legata ad alcune assunzioni sulla natura di

TAB. 1. I sistemi percettivi secondo Gibson

Sistema percettivo	Modalità di raccolta di informazioni	Raccoglie informazioni su
Orientamento di base	Non specifica, automatica, inconscia	Variazioni posturali rispetto alla gravità
Uditivo	Ascoltare	Natura e posizione di eventi esterni
Aptico	Toccare	Contatto con oggetti, loro forma, solidità e viscosità
Gusto/olfatto	Assaggiare/annusare	Valore nutritivo e biochimico di sostanze
Visivo	Guardare	Caratteristiche di oggetti animati e inanimati, superfici di supporto, eventi, luoghi

certi processi mentali. Per fare solo un esempio, Gibson propose una concezione unitaria del sistema visivo, mentre (come vedremo) esistono oggi numerosi risultati empirici che sembrano suggerire l'esistenza di sistemi multisensoriali, che coinvolgono meccanismi visivi diversi, per rilevare proprietà di oggetti o rappresentare lo spazio. Più importante per i nostri fini in questo volume è invece l'enfasi di Gibson sull'aspetto che diventa cruciale quando si adotta una prospettiva multisensoriale, ossia il fatto che i sistemi percettivi cooperano fra loro in molte maniere diverse. In molti casi raccolgono informazioni sulle medesime proprietà ambientali, anche se attraverso canali sensoriali diversi. L'informazione disponibile dall'ambiente in questo caso è ridondante, e la cooperazione ha l'effetto di rendere la percezione più robusta e affidabile [Ernst e Bülthoff 2004]. In altri casi l'informazione raccolta riguarda proprietà diverse dello stesso oggetto. L'informazione in questo caso è complementare, e la cooperazione rende la percezione più ricca e completa. Ad esempio, quando manipoliamo un oggetto che stiamo guardando, la mano può dare informazioni sulla parte posteriore anche quando questa non è visibile [Newell *et al.* 2001]. In altri casi ancora, le informazioni raccolte dai diversi canali sensoriali riguardano le stesse proprietà ma sono in contraddizione fra loro. In quest'ultimo caso, la cooperazione fra i diversi sistemi può dar luogo a interpretazioni percettive inaspettate, che costituiscono veri e propri percetti «emergenti», non riducibili ai canali sensoriali coinvolti o alla loro somma.

4. METODI PER LO STUDIO DELLA PERCEZIONE MULTISENSORIALE

Prima di addentrarci nello studio della multisensorialità, è utile anticipare alcuni concetti che riguardano le metodologie di studio che possono essere utilizzate per affrontare problemi di percezione multisensoriale. Come abbiamo già detto, lo studio del funzionamento della mente si basa oggi sulla combinazione di approcci metodologici diversi ma complementari. Da questo punto di vista la ricerca sui processi multisensoriali non fa eccezione. Da un lato continuano a essere di fondamentale importanza i metodi della *psicologia sperimentale* e della *psicofisica*, nei quali vengono misurate risposte comportamentali quali l'accuratezza, la precisione, o i tempi di risposta nello svolgimento di un compito, o ancora i parametri cinematici associati ai movimenti richiesti dal compito stesso. Contemporaneamente, lo studio della mente si è arricchito negli ultimi decenni di molteplici strumenti in grado di sondare o influenzare il funzionamento del cervello durante lo svolgimento dell'attività mentale richiesta dal compito. Fra queste nuove tecniche delle neuroscienze cognitive spiccano alcune metodiche di neuroimmagine quali la **risonanza magnetica funzionale** (fMRI, *functional Magnetic Resonance Imaging*) o la **tomografia a emissione di positroni** (PET, *Positron Emission Tomography*), che misurano il consumo metabolico di sostanze quali l'ossigeno o il glucosio nelle diverse aree cerebrali e consentono di inferire in maniera indiretta il coinvolgimento dell'area cerebrale attiva nel compito svolto. Oppure metodiche elettroencefalografiche come lo studio dei **potenziali elettrici evento-correlati** (ERP, *Event-Related Potentials*), che misurano in maniera non invasiva i potenziali elettrici generati dall'attività cerebrale durante lo svolgimento di un compito sperimentale grazie all'applicazione di elettrodi sul cuoio capelluto. O ancora metodiche di stimolazione della corteccia cerebrale quali la **stimolazione magnetica transcranica** (TMS, *Transcranial Magnetic Stimulation*), che si basano sul principio dell'induzione elettromagnetica per generare temporanee interferenze elettriche in aree più o meno circoscritte della corteccia cerebrale.

Tutti questi approcci vengono impiegati sia con individui neurologicamente sani, sia con pazienti con disturbi evolutivi o acquisiti del sistema nervoso centrale. Inoltre sono stati impiegati anche nello studio di individui con forme congenite o acquisite di deprivazione sensoriale grave (per esempio, nei ciechi o nei sordi) e più recentemente nello studio di individui portatori di protesi (come ad esempio gli impianti cocleari per l'udito o protesi che sostituiscono parti del corpo amputate) e perfino di individui con arti trapiantati. Infine lo studio della mente, nelle neuroscienze cognitive moderne e nello studio della multisensorialità in particolare, si avvale di modelli simulativi al computer quali, ad esempio, i modelli connessionisti di reti neurali artificiali. In questo volume ci limiteremo ad accennare alcuni degli aspetti fondamentali di ciascuna tecnica nel momento in cui descriveremo esperimenti specifici. Per una trattazione dettagliata delle diverse metodiche delle neuroscienze cognitive rimandiamo invece al testo di Bonfiglioli e Castiello [2005].

In generale, e in maniera indipendente dalla tecnica utilizzata, è possibile raggruppare le ricerche sulla multisensorialità in tre categorie generali. Fra queste, la prima comprende le ricerche che hanno esaminato il «valore aggiunto» della multisensorialità rispetto a situazioni di stimolazione unisensoriale. La seconda, le ricerche che hanno esaminato i contesti all'interno dei quali avviene l'integrazione multisensoriale e, viceversa, i contesti nei quali l'integrazione multisensoriale non si verifica nonostante la presenza di molteplici stimolazioni sensoriali. La terza, infine, comprende le ricerche che hanno esaminato il contributo e le peculiarità di ciascuna modalità sensoriale quando i diversi canali sensoriali forniscono informazioni fra loro discordanti. Di seguito tratteremo ciascuno di questi approcci più nel dettaglio, fornendo alcuni esempi paradigmatici.

Nelle ricerche raggruppabili nella prima categoria l'obiettivo è il confronto fra stimolazioni unisensoriali e multisensoriali. Questo richiede la creazione di contesti sperimentali nei quali sia possibile stimolare un singolo canale sensoriale separatamente dagli altri. La cosa interessante è che, proprio perché la nostra esperienza quotidiana è in maniera preponderante multisensoriale, le situazioni artificiose nelle quali riceviamo stimolazioni unisensoriali possono dare luogo a percezioni insolite. Un gioco ripetutamente proposto alle feste di adulti e bambini si presta bene a illustrare la condizione di confronto fra unisensoriale e multisensoriale, e la sorpresa che proviamo di fronte a una stimolazione artificiosamente unisensoriale. Si tratta di chiedere a un amico di riconoscere il cibo che gli viene proposto (per esempio una fragola o un pezzo di banana) sulla base delle sole informazioni gustative, e in assenza di indizi visivi, olfattivi e acustici (al partecipante viene messa una benda sugli occhi, e vengono tappati il naso e le orecchie). Lo stupore del partecipante (e il divertimento di coloro che scelgono e propongono i diversi cibi) deriva dalla difficoltà, spesso accompagnata da errori grossolani, nell'individuare il cibo proposto. Come vedremo nel capitolo 6, l'esperienza del sapore dei cibi è primariamente un'esperienza multisensoriale. Per quanto ciascuno di noi abbia la sensazione che il sapore dei cibi abbia origine sulla lingua e sul palato, in realtà la percezione del sapore dipende prevalen-

temente dalle informazioni olfattive che dal palato raggiungono la cavità nasale, e può essere influenzata dalle informazioni visive circa la forma e il colore del cibo appena ingerito, dalle sensazioni tattili e termiche nel cavo orale e persino dalle informazioni acustiche che derivano dalla masticazione.

Questo primo esempio evidenzia quanto l'esperienza multisensoriale possa essere pervasiva nella nostra interazione con il mondo esterno anche quando non ce ne rendiamo conto. Un ulteriore esempio può essere utile per illustrare come il confronto fra condizioni unisensoriali e multisensoriali possa rivelare il valore aggiunto della multisensorialità nella percezione ordinaria. Tutti noi siamo consapevoli di quanto sia difficile seguire una conversazione in un contesto rumoroso, in cui molteplici conversazioni si accavallano le une alle altre. Siamo però anche consapevoli che la comprensione del messaggio verbale che ci interessa ci risulta più semplice quando possiamo vedere i movimenti delle labbra del parlante. Questo beneficio dato dall'aggiunta dell'informazione visiva non può essere attribuito alla sola capacità di leggere il messaggio sulle labbra del parlante (abilità peraltro ben poco sviluppata nelle persone dotate di un udito nella norma), ma dipende piuttosto dall'integrazione che il cervello effettua fra le informazioni acustiche (messaggio verbale) e le informazioni visive (movimento delle labbra del parlante). L'integrazione, come vedremo nel resto di questo volume, può risultare particolarmente preziosa quando almeno uno dei canali sensoriali sta fornendo informazioni degradate o viziate dal cosiddetto «rumore», cioè dall'attività di disturbo che può essere presente nel canale che convoglia l'informazione (nella teoria dell'informazione si parla appunto di rapporto segnale/rumore).

Questo esempio di interazione fra messaggio percepito attraverso l'udito e i movimenti delle labbra, percepiti attraverso la visione, ci consente di introdurre il secondo tipo di approccio allo studio della multisensorialità, ovvero lo studio dei contesti di stimolazione multisensoriale in cui si verifica integrazione delle informazioni provenienti dai diversi organi di senso. Il problema teorico sottostante è il seguente: se l'ambiente nel quale viviamo è *sempre* multisensoriale, come fa il nostro sistema percettivo a scegliere quali informazioni multisensoriali integrare e quali invece non integrare? Ad esempio, quando conversiamo in un contesto in cui ci sono molte altre conversazioni attorno a noi, come fa il sistema a scegliere quale messaggio verbale vada integrato con le informazioni visive circa i movimenti delle labbra del parlante? Una gran quantità di studi sulla percezione multisensoriale è stata ideata per manipolare in maniera sistematica il grado di sovrapposizione temporale e spaziale tra le informazioni provenienti dai diversi canali sensoriali e cercare di fornire una risposta a tale quesito. In estrema sintesi, la risposta è che il nostro sistema percettivo tende a considerare come provenienti dalla medesima fonte (e quindi integrare) le informazioni sensoriali che originano da posizioni dello spazio *pressoché* identiche, e che si verificano *pressoché* nel medesimo istante temporale. In una prospettiva evolutiva, si potrebbe affermare che il sistema percettivo si è adattato a un ambiente in cui informazioni concomitanti nello spazio e nel tempo originano dal medesimo evento ambientale, ma applica questa regola con un «intervallo di tolleranza».

In molti contesti della vita quotidiana questo intervallo di tolleranza è trascurabile. Se un parlante è immediatamente di fronte a noi il suono della sua voce proverrà esattamente dal punto in cui vediamo muoversi le sue labbra, e le due informazioni saranno perfettamente sincrone nel tempo (se non consideriamo il ritardo di qualche decina di millesimi di secondo che l'informazione visiva ha rispetto a quella acustica, per via della più rapida trasmissione neurale di quest'ultima). Ma che dire dei contesti di comunicazione a distanza? Nelle teleconferenze, o più semplicemente nelle videocomunicazioni che tutti noi possiamo facilmente attivare dal nostro computer, il parlato e i movimenti delle labbra sono spesso asincroni, e i ritardi possono essere di diverse centinaia di millisecondi. Inoltre, nelle videocomunicazioni o al cinema il messaggio parlato non proviene dalla posizione dello schermo dove vediamo muoversi le labbra del nostro interlocutore, ma origina da altoparlanti più o meno lontani da quel punto. Eppure il nostro sistema percettivo sembra essere in grado di accettare discrepanze temporali di centinaia di millisecondi [Spence e Squire 2003] e discrepanze spaziali anche cospicue prima di rinunciare all'unità percettiva dell'oggetto multisensoriale.

Questa tenacia nel ricercare l'unità percettiva delle informazioni multisensoriali pressoché corrispondenti nello spazio e nel tempo è anche all'origine delle percezioni illusorie che possono sorgere quando le informazioni provenienti dai diversi organi di senso sono discordanti. Questo terzo tipo di approccio alla ricerca multisensoriale è esemplificato molto chiaramente dal cosiddetto **effetto McGurk**: un'illusione visuoacustica, nuovamente legata all'interazione fra messaggio parlato e movimenti percepiti delle labbra, descritta per la prima volta da Harry McGurk e John MacDonald nel 1976. Nello studio originale i partecipanti all'esperimento osservavano un video che mostrava il volto di una giovane donna mentre pronunciava la sillaba «ga», ma udivano invece il suono della sillaba «ba». In questo contesto di stimolazione visuoacustica discrepante i partecipanti non abbandonavano l'unità percettiva delle due informazioni, ovvero associavano la sillaba udita al movimento delle labbra osservato, ma riportavano sistematicamente di aver udito la sillaba «da» (in realtà mai presentata!). È importante notare che in questa situazione è sufficiente chiudere gli occhi per

percepire correttamente la sillaba effettivamente pronunciata (ovvero «ba»). Inoltre, non si commette alcun errore quando il video viene doppiato in modo che le informazioni visuoacustiche siano corrispondenti [McGurk e MacDonald 1976]. Dunque, è precisamente l'interazione multisensoriale a generare il percetto illusorio. Avremo modo di riparlarne più avanti.

Le situazioni di conflitto fra informazioni presenti in diversi aspetti dello stimolo vengono spesso utilizzate dagli psicologi sperimentali come mezzo per comprendere il funzionamento dei sistemi percettivi e, più in generale, della mente (vedi ad esempio Gregory [1989]; o il classico lavoro di Stroop [1935]). Nel caso della multisensorialità, lo studio dei contesti nei quali l'unità percettiva integra fra loro informazioni discrepanti è stato e continua a essere un approccio di grande utilità per comprendere i rapporti di dominanza fra le modalità sensoriali e i criteri in base ai quali questa dominanza può essere modificata. Come vedremo in maniera più approfondita nel resto di questo volume, nell'integrare le informazioni provenienti dai diversi organi di senso il cervello tende a basarsi di volta in volta sul canale sensoriale che offre l'informazione più affidabile. In questo senso la multisensorialità si caratterizza chiaramente come un aspetto adattivo del nostro sistema cognitivo, in grado di ottimizzare le nostre percezioni rispetto alle molteplici condizioni di rumore e incertezza che caratterizzano l'ambiente nel quale viviamo.

5. COME È FATTO QUESTO LIBRO

I fenomeni multisensoriali di cui ci occupiamo in questo volume sono raggruppati in quattro grandi categorie, a ognuna delle quali è dedicata una parte suddivisa a sua volta in capitoli che ne approfondiscono aspetti particolari.

Nella prima (capp. 1-3) ci occupiamo della percezione del *corpo*. Ci proponiamo di presentare lo stato delle conoscenze attuali sul ruolo della multisensorialità nei processi percettivi e motori che ci consentono di sapere quali siano le parti del nostro corpo e come sono fatte, quali posizioni occupano nello spazio, come le stiamo muovendo.

A seguire (seconda parte, capp. 4-7) prendiamo invece in esame la percezione multisensoriale degli *oggetti*. Anche il nostro corpo, naturalmente, può essere considerato un oggetto percettivo, e come vedremo la distinzione fra una parte del proprio corpo e un oggetto esterno non è, dal punto di vista dello studio della percezione, del tutto priva di problemi. Tanto è vero che anche un oggetto inanimato può in certe condizioni venire incorporato nella nostra rappresentazione

corporea. Ma la distinzione fra corpo e oggetti, che implicitamente proponiamo nella struttura del libro, appare giustificata dal tipo di ricerche che hanno caratterizzato queste due aree: infatti lo studio del riconoscimento e della percezione di oggetti si è in larga parte concentrato sugli «oggetti» nel senso di artefatti con specifiche funzioni, ossia gli oggetti della nostra vita quotidiana, solidi, con una precisa struttura tridimensionale. Fa eccezione, rispetto a questa caratterizzazione, una tipologia di oggetti particolarmente interessante in una prospettiva multisensoriale, ossia i cibi. Le sostanze di cui ci cibiamo possono essere considerate artefatti – pensiamo al piatto che ci viene presentato al ristorante – ma possono anche essere semplici oggetti naturali – pensiamo a un frutto o a un tubero. In ogni caso, i cibi hanno un valore biologico fondamentale e per questo motivo non è sorprendente che i processi percettivi coinvolti nella nutrizione siano fra quelli in cui l'interazione multisensoriale è più intensa. Il capitolo 6 della seconda parte è dedicato esclusivamente al cibo.

Il tema della struttura tridimensionale degli oggetti si collega al tema che segue (terza parte, capp. 8-11), quello della percezione multisensoriale dello *spazio*. Lo spazio è un tema che ha preoccupato la filosofia per secoli, a partire da Democrito e Aristotele fino a Cartesio e Kant, oltre alla psicologia della percezione e alle neuroscienze cognitive moderne. Ci proponiamo di mostrare come le ricerche recenti in psicologia e neuroscienze suggeriscano una distinzione fra diversi tipi di spazio, distinguibili sulla base dei diversi tipi di interazioni multisensoriali coinvolte nella percezione e nell'azione.

Nella quarta parte (capp. 12-14), ci occupiamo della percezione multisensoriale del *tempo*, o più precisamente delle proprietà temporali degli eventi, come la *sincronicità*, la *durata* e il *ritmo*.

Nell'epilogo, infine, proviamo a tirare le somme di tutto il nostro discorso sulla multisensorialità. Questo capitolo conclusivo prova dunque a riassumere quelle che sembrano essere le risposte alle domande molto generali poste all'inizio di questa introduzione, e conclude discutendo brevemente alcuni problemi teorici che riguardano i modelli generali della percezione multisensoriale. Si tratta di tematiche che sono sicuramente destinate a essere oggetto di ulteriore discussione in futuro. Qui speriamo perlomeno di essere riusciti a illustrare i principi multisensoriali che regolano la percezione del corpo, degli oggetti, dello spazio e del tempo, e a dare al lettore un'idea di come questi trovino applicazione anche a molti aspetti della comunicazione umana, da quella più ovvia che coinvolge il linguaggio, fino a molte altre problematiche in cui entrano in gioco le relazioni sociali e/o le tecnologie per interfacciare una macchina a un operatore umano.

PARTE PRIMA

Il corpo

CAPITOLO 1

Il corpo come oggetto multisensoriale

La percezione del nostro corpo non è dovuta solo alle esperienze derivanti dai sensi corporei come il tatto e la propriocezione, ma è invece il risultato dell' integrazione fra diverse informazioni sensoriali. Le molteplici rappresentazioni che possediamo del nostro corpo sono di natura multisensoriale, e hanno la funzione di renderci consapevoli di come siamo fatti, delle posizioni relative dei nostri arti e di come possiamo muoverli o li stiamo effettivamente muovendo. Presumibilmente, dall'insieme dei processi multisensoriali relativi alla percezione del corpo trae origine il vissuto che il corpo che percepiamo ci appartenga e sia la sede del nostro Sé.

1. LA PERCEZIONE DEL CORPO

In un classico della psicologia della percezione [von Fieandt 1966], il capitolo sulla percezione del Sé si apre invitando il lettore a fare alcuni esperimenti usando un semplice specchio da casa. Provate anche voi. Di fronte allo specchio, ruotate la mano in modo da portare il palmo verso il corpo, sollevatela fino all'altezza del viso, e iniziate a picchiettare con il dito indice sulla vostra fronte. Fate questo alternando periodi in cui vi guardate nello specchio e periodi in cui tenete gli occhi chiusi. Avete notato una differenza? Durante il tocco a occhi chiusi l'esperienza è coerente: la sensazione è localizzata sulla fronte e il movimento del braccio è nella direzione del volto, grazie alla flessione del braccio che avvicina la mano al corpo. Ma quando si aprono gli occhi le cose cambiano. Ora l'esperienza non è più coerente, diventa quasi paradossale. La sensazione infatti

non è più localizzata sulla propria fronte, ma sulla fronte del corpo riflesso, ossia nello spazio virtuale che sta percettivamente dietro la superficie dello specchio e non sul *proprio* corpo. Se non ci avete fatto caso, provate ancora una volta. Confrontate la sensazione con gli occhi chiusi e aperti, chiedendovi: «dove sento il tocco?» e facendo attenzione alla differenza. Naturalmente, a occhi aperti l'esperienza è in parte paradossale, perché il movimento del braccio in flessione, che pure viene sentito come tale, in un certo senso sembra anche allontanare la mano dal corpo, portandola verso il corpo riflesso. Nella sensazione muscolare il movimento rimane una flessione, ma nel percetto visivo è una estensione. Inoltre si ha l'impressione di avere ruotato il polso portando il palmo verso il corpo, ma in un certo senso quando si hanno gli occhi aperti sembra che il polso sia ruotato in modo da avere il palmo verso il corpo riflesso e il dorso, non il palmo, verso il proprio corpo.

Per von Fieandt, esperienze come questa dimostravano come la rappresentazione percettiva del corpo (da lui chiamata **immagine corporea fenomenica**) dipenda da un complesso di sensazioni, in cui quelle visive hanno un ruolo particolare per la tendenza a «visualizzare» il materiale sensoriale trasmesso da altri canali. Cosa intendesse allora von Fieandt con il termine «visualizzare» non è del tutto chiaro. Come vedremo in questa parte del libro, un gran numero di ricerche più recenti suggerisce che una possibile reinterpretazione potrebbe essere questa: il corpo è un oggetto multisensoriale e come tale la sua rappresentazione dipende da processi in cui i diversi sistemi sensoriali interagiscono fra di loro in maniera sofisticata e sorprendente. Inoltre, la rappresentazione del corpo è flessibile, estremamente sensibile ai mutamenti dei segnali sensoriali in ingresso, e assai poco influenzabile da quello che sappiamo razionalmente su come siamo fatti. Al punto da darci l'impressione di poter toccare il nostro corpo nello specchio, come in una specie di esperienza extracorporea simile a quelle che descrivono i praticanti di meditazione trascendentale o i pazienti epilettici. Infine, il canale sensoriale visivo sembra avere, in queste interazioni multisensoriali, uno status privilegiato.

Dalla pubblicazione del libro di von Fieandt (la cui prima versione, in finlandese, è addirittura del 1951) molti altri ricercatori si sono occupati della rappresentazione multisensoriale del corpo. Proviamo a farlo anche noi, cominciando con una domanda: attraverso quali sensazioni conosciamo il nostro corpo? La prima risposta che viene in mente è che la conoscenza del corpo passa attraverso i cosiddetti sensi corporei, ossia dalle sensazioni che ci derivano dalla superficie della cute, dai muscoli, dagli organi interni e dal nostro corpo in generale. I sensi corporei comprendono le sensazioni tattili, di temperatura e di

dolore, ma anche il senso di posizione del corpo. Quest'ultimo, un senso corporeo quasi segreto che si basa sull'elaborazione di tutte le informazioni sensoriali che derivano da muscoli, tendini e articolazioni per controllare e adattare continuamente la posizione, il tono e il movimento del nostro corpo, è stato definito **propriocezione** dal neurofisiologo inglese Charles Sherrington.

Sherrington coniò questo termine per differenziare la percezione del corpo dalla percezione degli oggetti esterni («esterocezione») e dei segnali interni dell'organismo («interocezione»), ma soprattutto per sottolinearne il ruolo indispensabile per avvertire il corpo come nostro. Tuttavia, parte della conoscenza del vostro corpo dipende anche dalle informazioni che vi restituisce la vista. Pensate a tutte le occasioni in cui avete avuto occasione di osservare il vostro corpo in maniera diretta, semplicemente abbassando gli occhi o portando lo sguardo verso le mani per guardarle mentre afferravano un oggetto. O pensate alle volte in cui avete visto il vostro corpo in maniera indiretta, riflesso negli specchi di casa o nelle vetrine dei negozi mentre camminate, raffigurato in una postura statica all'interno di una fotografia, mostrato nel dinamismo di un'azione in una videoripresa, o come una sagoma scura quando le condizioni di illuminazione ne proiettano l'ombra. Proseguendo in questa ricerca delle sensazioni che consentono di conoscere il corpo potreste anche notare che il vostro corpo può essere ascoltato: mentre interagisce con l'ambiente, mentre scricchiola nelle giunture, mentre si lamenta per la fame, o quando emettete dei suoni spingendo l'aria attraverso le corde vocali. E può essere annusato: la fiorente industria dei profumi esiste per dare un odore al corpo, sovrastando gli odori che il corpo naturalmente produce in quanto oggetto biologico. Infine il corpo può essere gustato: tutti noi abbiamo in mente la sensazione di quando, nel gesto istintivo di alleviare il dolore di una piccola ferita a un dito, portiamo il dito alla bocca e sentiamo il sapore della nostra pelle e del nostro sangue.

Ma basta tutto questo per dire che il corpo è un oggetto multisensoriale? Non necessariamente. Benché si possano raccogliere informazioni sul corpo da molteplici fonti di informazione, la nostra conoscenza del corpo potrebbe basarsi in ultima istanza sulle sole sensazioni corporee. Dopotutto le sensazioni corporee sono *uniche* e *private*: di nessun altro oggetto o corpo che non sia il nostro possiamo avere, a oggi, una completa esperienza somatosensoriale. In secondo luogo le sensazioni corporee sono praticamente *sempre presenti*: per usare le parole del filosofo Merleau-Ponty [1945], il nostro corpo è il solo oggetto che non ci abbandona mai. Proprio le sensazioni corporee potrebbero essere quell'aspetto elementare e imprescindibile dal quale dipende come percepiamo il nostro corpo. Infine, le sensazioni corporee sono codificate sotto forma di

mappa ordinata in una specifica zona del cervello. Sembra naturale pensare che questa mappa rappresenti la base della percezione del corpo. Invece, come vedremo fra breve, parlare del corpo come oggetto multisensoriale in maniera non ingenua significa anzitutto dimostrare che il corpo percepito non può essere ridotto alle sole afferenze somatosensoriali e alla loro rappresentazione corticale.

2. NON SOLO AFFERENZE SOMATOSENSORIALI

Come già anticipato nell'introduzione, le sensazioni che provengono dalla cute, dai muscoli, dalle articolazioni e dagli altri recettori presenti nei tessuti corporei vengono convogliate in una zona della corteccia nota come **corteccia somatosensoriale primaria** (o S1), nel giro postcentrale. L'immagine raffigurata in figura 1.1 A) mostra questa disposizione ordinata delle specializzazioni dei territori corticali in S1. Esplorando l'immagine da destra verso sinistra (ovvero, dai territori corticali più mediali a quelli più laterali, in questa visione coronale di un singolo emisfero cerebrale) si incontrano popolazioni di neuroni che, in maniera ordinata, ricevono segnali dai genitali, dalle dita del piede, dal piede, dalla gamba, e via via a salire fino alla spalla, il braccio, la mano e le dita della mano. Adiacente al territorio corticale del pollice, troviamo poi le afferenze provenienti dagli occhi, dal naso e dal volto in generale, che a loro volta sfumano nel territorio corticale del cavo orale, della lingua, della faringe e delle sensazioni intraddominali. Questa mappa ordinata delle afferenze somatosensoriali è stata ipotizzata per la prima volta da John Hughlings-Jackson nel 1863, osservando la sequenza delle sensazioni tattili riportate da alcuni pazienti epilettici dopo la crisi. Successivamente, è stata documentata da studi di stimolazione elettrica nell'animale e nell'uomo [Penfield e Boldrey 1937], ed è stata descritta anche attraverso l'utilizzo della risonanza magnetica funzionale [Hlustík *et al.* 2001].

Un primo aspetto che appare evidente è che questa mappa riduce la tridimensionalità del corpo a una sola dimensione, ovvero tutto il corpo è rappresentato su una sola «striscia» di corteccia cerebrale. Come conseguenza di ciò alcuni distretti corporei fra loro distanti finiscono con l'occupare territori corticali vicini (vedi per esempio quanto accade per il volto e per la mano che pur essendo distanti sul nostro corpo sono invece adiacenti sulla mappa cerebrale). La prima conclusione che possiamo quindi trarre è che se la mente dovesse rappresentare il corpo basandosi esclusivamente su questa mappa, la rappresentazione corrispondente includerebbe inevitabilmente una distorsione delle relazioni spaziali fra le diverse parti del corpo.

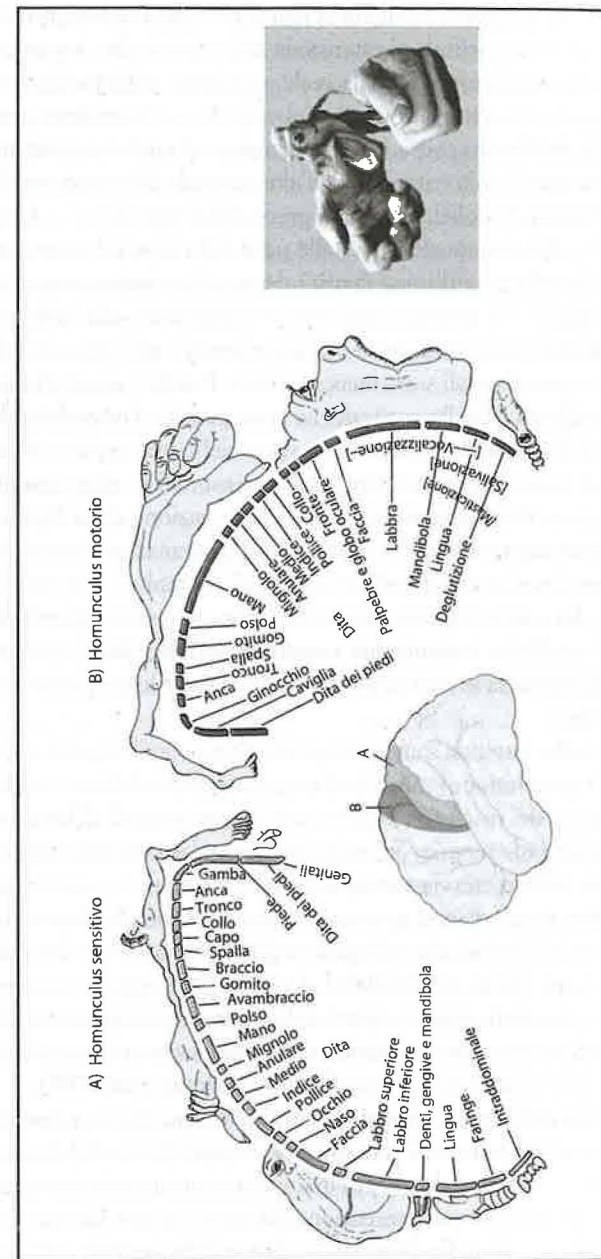



fig. 1.1. A sinistra, la mappa del corpo descritta da Penfield nell'area somatosensoriale primaria (indicata con il grigio chiaro e la lettera A sul disegno del cervello). Quando viene stimolata una zona della corteccia somatosensoriale, il soggetto riferisce di sentire una sensazione tattile nella parte corrispondente del corpo (vedi disegno dell'homunculus sensitivo). Nell'area grigio scura (indicata con la lettera B sul disegno del cervello) si trova l'area motoria primaria e una mappa corporea simile a quella somatosensoriale (ma non identica). Quando questa seconda area viene stimolata si ottengono attivazioni muscolari in specifici distretti corporei. A destra, l'homunculus corticale: come dovrebbe essere il nostro corpo se le dimensioni relative degli arti fossero pari all'area relativa di corteccia a esse dedicate.

Un secondo aspetto altrettanto evidente in figura 1.1 è che i diversi distretti corporei occupano territori corticali di estensione differente e che, soprattutto, non corrispondono necessariamente alla reale estensione della parte corrispondente. Ad esempio, il territorio corticale del volto ha un'estensione quasi paragonabile a quella dell'intero corpo (esclusa la mano e il piede), nonostante esso sia di fatto molto più piccolo (circa 1/8 dell'altezza totale del corpo, stando agli insegnamenti classici di Policleto, artista greco del V secolo a.C.). Questa discrepanza fra la dimensione effettiva delle parti del corpo e l'estensione del territorio corticale riflette la diversa densità dei recettori somatosensoriali nei vari distretti corporei. Ad esempio, nel volto e soprattutto sulle labbra vi sono una gran quantità di recettori, mentre nell'avambraccio, nel tronco o nelle gambe i recettori somatosensoriali sono meno presenti. Poiché i recettori hanno una proiezione ordinata a livello corticale, ne consegue che l'estensione del territorio corticale dedicato a una determinata parte del corpo è più o meno ampio in funzione di quanti sono i recettori somatosensoriali in essa presenti. La seconda conclusione che possiamo trarre dall'osservazione della figura è dunque che se la nostra mente dovesse ricavare una rappresentazione basandosi sulla sola mappa somatosensoriale dovrebbe avere un'idea molto distorta delle dimensioni relative del corpo. Questa distorsione viene spesso rappresentata con il modello del cosiddetto **homunculus somatosensoriale**, che ci consente di visualizzare la differenza fra le reali proporzioni del corpo e le proporzioni di corteccia corrispondenti.

Infine, a livello della corteccia somatosensoriale nemmeno il lato del corpo dal quale proviene la stimolazione sensoriale è necessariamente definito. Infatti, benché la maggior parte dei recettori corporei invii i propri segnali all'emisfero cerebrale controlaterale (ad esempio, un tocco alla mano destra viene anzitutto registrato da neuroni della corteccia somatosensoriale sinistra), le cortecce somatosensoriali dei due emisferi sono in costante comunicazione fra loro attraverso fasci di fibre dette callosali, che collegano regioni omologhe nei due emisferi cerebrali [Iwamura, Taoka e Iriki 2001]. Dunque potremmo concludere che, basandosi sulle sole informazioni presenti nella corteccia somatosensoriale, la nostra mente potrebbe non essere nemmeno in grado di sapere con certezza da quale lato del corpo provengano le sensazioni tattili [Braun *et al.* 2005].

Queste **distorsioni della rappresentazione corporea** legate alla rappresentazione del corpo disponibile nella corteccia somatosensoriale primaria (distanze relative, dimensioni relative e lato del corpo) in una certa misura esistono e possono produrre illusioni tangibili nella percezione del nostro corpo. Un esempio classico riguarda la percezione delle dimensioni relative delle parti del corpo ed

è stato descritto più di 150 anni fa dal fisiologo e anatomista tedesco Ernst Heinrich Weber. Nel 1834, ragionando sull'origine delle sensazioni e delle percezioni corporee, egli si accorse di un'illusione tattile che si verificava quando cercava di confrontare la distanza fra coppie di punti applicati sulla cute. Weber notò che la distanza fra i due stimoli tattili sembrava maggiore quando questi erano somministrati a regioni della cute con alta densità di recettori somatosensoriali (per esempio, il dito indice), rispetto a quando gli stessi stimoli erano somministrati a regioni di cute con bassa densità di recettori (per esempio, l'avambraccio). Potete sperimentare facilmente voi stessi questa illusione toccando con due matite appuntite due punti distanti pochi centimetri sul dito indice e poi applicando la stessa stimolazione sull'avambraccio. Quello che dovreste notare è che, nonostante abbiate mantenuto costante la distanza fra le punte della matita, le punte sembrano più distanti sul dito indice rispetto all'avambraccio. Questa illusione, che naturalmente funziona ancora meglio se a darvi la stimolazione tattile è una seconda persona e voi non avete alcuna idea circa l'effettiva distanza fra le due punte di matita, è oggi nota col nome di **illusione di Weber**. Essa riflette il fatto che nello stimare la distanza fra due punti sulla cute la mente sembra utilizzare, almeno in parte, la metrica distorta disponibile al livello della corteccia somatosensoriale, nella quale il dito indice ha un'estensione molto maggiore rispetto a quella dell'avambraccio, e rivela che l'anomala rappresentazione del corpo disponibile nella corteccia somatosensoriale è in una certa misura in grado di influenzare alcune delle nostre percezioni corporee.

Eppure, nella vita di tutti i giorni, non crediamo che il dito indice occupi una superficie corporea più estesa dell'avambraccio, né tanto meno confondiamo la mano con il volto, o la stimolazione sul lato destro del corpo con quella sul lato sinistro. Come mai? Evidentemente partendo dalle informazioni presenti nella corteccia somatosensoriale la nostra mente è in grado di formare una rappresentazione del corpo di più alto livello, che meglio corrisponde alla reale struttura del soma e alla sua vera metrica. L'idea attualmente più accreditata è che la **geometria del corpo** (ovvero la definizione delle sue proprietà metriche, come la lunghezza o l'estensione), la **mereologia del corpo** (ovvero la sua differenziazione in parti categorialmente distinte) e la **topologia del corpo** (ovvero la specificazione delle relazioni spaziali fra le parti stesse) siano il prodotto di un'integrazione di informazioni multisensoriali e motorie e in questo senso non si possa ridurre il corpo alle sole sensazioni che ricaviamo dai sensi corporei. Due esempi possono aiutare a chiarire questo concetto.

Il primo esempio è tratto da una ricerca condotta nel 2004 da Taylor-Clarke, Jacobsen e Haggard utilizzando proprio l'illusione di Weber. Come nell'esem-

pio descritto in precedenza i partecipanti a questo studio dovevano giudicare quale di due distanze tattili somministrate sul dito indice e sull'avambraccio era maggiore. A parità di distanze tattili, i soggetti tendevano a giudicare la distanza sul dito indice maggiore rispetto a quella sull'avambraccio, in accordo con il principio sottostante l'illusione di Weber. Tuttavia, Patrick Haggard e colleghi introdussero una manipolazione tanto semplice quanto ingegnosa: dopo una prima misura dell'illusione di Weber chiesero ai soggetti di osservare per circa un'ora un'immagine del loro avambraccio distorto da una lente di ingrandimento, in modo da apparire più esteso. Dopo questa prolungata esposizione a una distorsione (visiva) del loro corpo, i soggetti dell'esperimento cominciavano a giudicare sistematicamente le distanze tattili su questa parte del corpo come più ampie. Al contrario, se la distorsione visiva rendeva l'avambraccio più piccolo i soggetti giudicavano le distanze tattili come più ridotte [Taylor-Clarke, Jacobsen e Haggard 2004]. Questi risultati dimostrano che il nostro cervello opera una costante trasformazione dell'informazione tattile originariamente disponibile in corteccia somatosensoriale, adattandola a quella che ritiene essere la forma e la metrica del corpo così come appare su base visiva.

Un secondo esempio che può aiutare a comprendere come la metrica del nostro corpo non si limiti a ciò che la nostra mente può leggere nella mappa somatosensoriale è la cosiddetta **illusione di Pinocchio**. L'illusione di Pinocchio, descritta per la prima volta da James Lackner nel 1988, è la sorprendente variante di un'illusione propriocettiva causata dalla vibrazione del tendine del bicipite. Se chiedete a una persona di tenere il braccio destro semiflesso e gli occhi chiusi, una vibrazione continua applicata al tendine del bicipite produrrà la sensazione cinestesica che l'articolazione del braccio si stia estendendo. Questo curioso fenomeno avviene poiché la vibrazione stimola dei recettori muscolari del bicipite (detti fusi neuromuscolari) che normalmente sono attivi durante lo stiramento del muscolo. Lackner [1988] osservò che se ai partecipanti veniva chiesto di toccare la punta del naso con il dito indice del braccio semiflesso, la vibrazione non causava solo l'illusione di un allontanamento dell'avambraccio dal corpo ma anche la sensazione di un allungamento del naso del partecipante (da cui il nome di «illusione di Pinocchio»). La co-occorrenza dello spostamento dell'avambraccio nello spazio e la persistente sensazione di contatto fra la punta del dito e il naso possono quindi produrre una drastica modulazione della metrica del corpo, con scarsa considerazione per le conoscenze precedenti dell'individuo sulla forma del suo naso, ma in virtù dell'integrazione delle molteplici informazioni somatosensoriali che ricaviamo dai muscoli.

3. OLTRE LE AFFERENZE SOMATOSENSORIALI: SENTIRE IL CORPO IN UN ARTO FINTO

A questo punto dovrebbe essere evidente come la percezione del nostro corpo non si basi sulle sole informazioni presenti nella corteccia somatosensoriale primaria, e in generale sulle sole informazioni provenienti dai sensi corporei. Ad esempio, l'esposizione ripetuta a un'immagine distorta del proprio corpo può modulare la metrica percepita di quello stesso distretto corporeo [Taylor-Clarke, Jacobsen e Haggard 2004], suggerendo un ruolo della visione sulla costruzione della rappresentazione geometrica del corpo. Ma quanto avanti possiamo spingerci in questa concezione del corpo come oggetto multisensoriale e quanto veramente determinante è il ruolo della visione nella costruzione della rappresentazione corporea? Alcune risposte sorprendenti a questa domanda arrivano dalle ricerche condotte nell'ultimo decennio sull'**illusione dell'arto finto**.

I primi resoconti sull'illusione dell'arto finto risalgono alla fine degli anni '30 del secolo scorso, quando Tastevin [1937] notò che un osservatore poteva in alcuni casi confondersi e giudicare come proprio un dito finto che spuntava da sotto un panno, quando il suo vero dito era nascosto a pochi centimetri di distanza. Diversi anni più tardi Welch [1972] notò che questo errore di attribuzione sembrava non permeabile alle conoscenze dell'osservatore; in altre parole, anche quando gli osservatori sapevano che il dito osservato apparteneva allo sperimentatore avevano comunque la sensazione che potesse essere proprio. L'interesse per questo tipo di illusione si è però definitivamente affermato nel 1998 quando Matthew Botvinick e Jonathan Cohen pubblicarono sulla rivista «Nature» una descrizione breve ma sistematica di questa illusione. Gli autori chiesero a un gruppo di soggetti di sedersi a un tavolo, appoggiando la loro mano sinistra sul piano e dietro a uno schermo verticale opaco che ne impediva la vista. Visibile di fronte al soggetto era invece posta una copia in gomma a dimensione reale di una mano sinistra con relativo avambraccio, e la consegna al soggetto era semplicemente quella di mantenere lo sguardo su questo arto finto. Nel frattempo lo sperimentatore stimolava con due pennelli la mano sinistra del soggetto e l'arto finto, cercando di sincronizzare quanto più possibile l'istante delle pennellate sulle due mani (vedi fig. 1.2).

Dopo 10 minuti di stimolazione i partecipanti riportavano la sensazione curiosa di percepire il tocco del pennello nella posizione in cui vedevano stimolata la mano finta. Inoltre, se venivano invitati a indicare la posizione della loro mano sinistra puntando sotto il tavolo con l'indice della mano destra, i partecipanti tendevano sistematicamente a indicare una posizione della loro

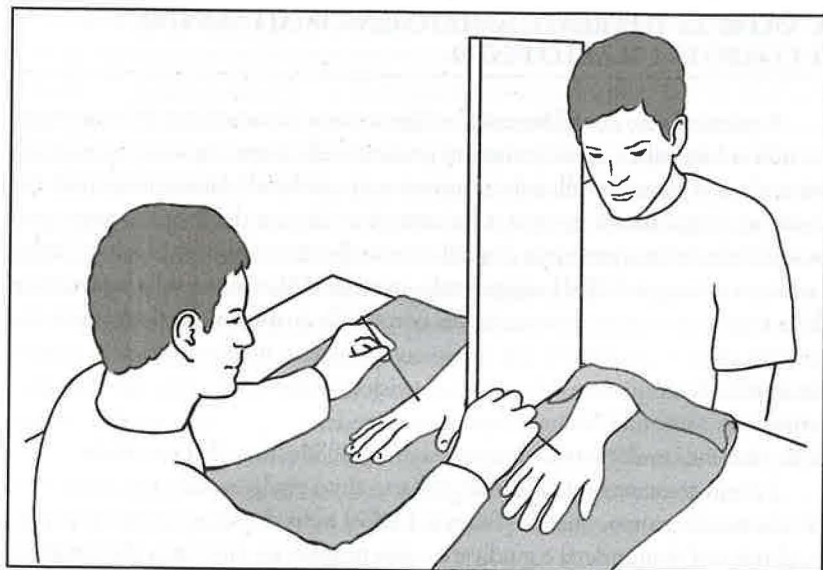



fig. 1.2. Nell'illusione dell'arto finto i soggetti ricevono uno stimolo tattile sulla propria mano nascosta mentre vedono una mano di gomma che viene toccata. Se lo stimolo visivo e quello tattile sono ben sincronizzati può insorgere l'impressione che la sensazione di tocco sia localizzata sulla mano di gomma.

vera mano sinistra spostata verso la posizione dell'arto finto, come se il loro senso di posizione per la mano sinistra fosse stato in qualche misura ricalibrato dalla ripetuta esposizione all'arto finto. Infine, la maggior parte dei partecipanti erano propensi a credere che la mano artificiale potesse effettivamente essere la loro mano [Botvinick e Cohen 1998]. Il semplice contesto sperimentale della mano finta, quindi, è in grado di evocare almeno tre illusioni distinte: anzitutto è in grado di provocare sensazioni tattili riferite a un arto finto (**cattura visiva del tatto** verso una posizione dalla quale non dovrebbe derivare alcuna sensazione tattile), in secondo luogo è in grado di provocare una ricalibrazione della posizione percepita del corpo nello spazio (**cattura visiva della propriocezione** verso una posizione nella quale non è presente alcun distretto del nostro corpo), in terzo luogo è in grado di influenzare ciò che il soggetto ritiene essere una possibile parte del proprio corpo. Come già aveva notato Welch [1972], e come vedremo più dettagliatamente in seguito, questa sensazione proprietaria verso una parte del corpo (solitamente descritta con il termine «ownership») sembra essere assai poco influenzata dalle conoscenze dell'osservatore, come risulta evidente dal fatto che tutti i partecipanti allo studio di Botvinick e Cohen [1998]

avevano ben chiaro da un punto di vista cognitivo che l'arto finto non era altro che una copia in gomma di un arto e che certamente non apparteneva al loro corpo. Tuttavia, alla domanda «Avevi la sensazione che l'arto di gomma fosse la tua mano?» la maggior parte di essi dava una risposta più che affermativa.

Negli ultimi dieci anni una gran quantità di studi ha cercato di definire le condizioni di insorgenza di questa illusione, cercando al contempo di offrirne una spiegazione (per una recente rassegna vedi Makin, Holmes e Ehrsson [2008]). Per quanto l'illusione possa emergere anche senza stimolazione simultanea della mano del soggetto e della mano finta (ovvero attraverso la semplice osservazione dell'arto finto, come nel caso descritto in origine da Tastevin [1937]; vedi anche Pavani, Spence e Driver [2000]), a oggi è assodato che essa tende a scomparire quando lo sperimentatore stimola le due mani in maniera asincrona [Tsakiris e Haggard 2005]. Questo fenomeno viene solitamente spiegato in base al fatto che la nostra mente tende ad attribuire un'origine comune agli eventi multisensoriali temporalmente sincroni (il cosiddetto **assunto di unitarietà**; Welch e Warren [1980]), mentre attribuisce origini distinte agli eventi multisensoriali temporalmente disgiunti. Dunque, un tocco alla propria mano e una pennellata su un arto finto che si verificano simultaneamente vengono attribuiti a un'origine comune, mentre un tocco e una pennellata che si verificano in due istanti diversi vengono attribuiti a due origini distinte. Inoltre, poiché la posizione esatta della mano finta nello spazio è molto ben identificabile (in quanto segnalata dalla modalità visiva), mentre la posizione della mano reale oltre lo schermo lo è solitamente molto di meno (in quanto segnalata dal solo senso di posizione), la nostra mente tende ad attribuire questo evento unico dove l'ha visto accadere e non dove l'ha percepito attraverso i sensi corporei, fidandosi quindi in misura maggiore del canale sensoriale in grado di fornire l'informazione più affidabile [Armel e Ramachandran 2003].

Se ci limitassimo a questa interpretazione del fenomeno basata sulla rilevazione di corrispondenze fra eventi sensoriali (un tocco e un evento visivo temporalmente sincrono), le implicazioni dell'illusione dell'arto finto sarebbero sorprendenti. Riprendendo la domanda che ci siamo posti all'inizio del paragrafo dovremmo concludere che è possibile percepire un *qualsiasi* oggetto dell'ambiente come parte del nostro corpo, purché si verifichino sensazioni tattili e visive temporalmente correlate. In altri termini, se al posto della mano finta vi fosse un blocchetto di legno i soggetti dovrebbero essere indotti a credere che quel pezzo di legno è parte del loro corpo, o che la loro mano si è trasformata in un parallelepipedo inerte di legno. Ma questo, in realtà, non accade. Per quanto sia possibile generare l'illusione dell'arto finto con versioni ben poco realistiche di

una mano (guanti da cucina riempiti di cotone, arti carnevaleschi o videoproiezioni bidimensionali, solo per fare alcuni esempi), l'illusione svanisce quando la mano finta viene ruotata di 90 gradi rispetto all'orientamento della mano del soggetto [Pavani, Spence e Driver 2000], quando viene eccessivamente rimpicciolita [Pavani e Zampini 2007], quando viene sostituita con una mano destra (mentre la mano nascosta è la sinistra; Tsakiris e Haggard [2005]), o quando, per riprendere l'esempio precedente, al posto della mano finta viene sistemato un parallelepipedo di legno [*ibidem*]. Questi ulteriori risultati mettono ben in chiaro che l'inganno della visione su ciò che può appartenere al nostro corpo è in realtà soggetto a vincoli, che presumibilmente hanno a che fare con ciò che la mente è disposta ad accettare come **immagine plausibile del corpo**. In altre parole, la costruzione multisensoriale della nostra rappresentazione del corpo sulla base delle sole informazioni in ingresso è limitata dalle rappresentazioni del corpo che la nostra mente *già* possiede.

4. LE MOLTEPLICI RAPPRESENTAZIONI DEL CORPO

A partire dal lavoro pionieristico dei neurologi inglesi Head e Holmes [1911], è stato suggerito che la nostra mente possieda almeno due rappresentazioni del corpo distinte. Da un lato, una rappresentazione largamente inconsapevole, basata su informazioni prevalentemente propriocettive e motorie, finalizzata soprattutto al controllo dell'azione. Dall'altro una rappresentazione più consapevole, largamente fondata su base visiva e finalizzata soprattutto alla nostra consapevolezza corporea. Diversi autori [Berlucchi e Aglioti 1997; Gallagher 1986; Paillard 1999] hanno associato al primo tipo di rappresentazione (originariamente denominata «schema posturale» da Head e Holmes [1911]) l'espressione **schema corporeo** (*body schema*; Schilder [1935]), mentre al secondo tipo di rappresentazione hanno associato l'espressione **immagine corporea** (*body image*; Head e Holmes [1911]; Lhermitte e Tchehrazi [1937]). Questa distinzione nasce principalmente da osservazioni in ambito neuropsicologico in merito al comportamento di pazienti che mostravano da un lato disturbi dello schema corporeo in assenza di problemi a carico dell'immagine corporea e dall'altro disturbi dell'immagine corporea in assenza di problemi di schema corporeo.

Un esempio del primo tipo di dissociazione (ovvero, danni a carico del solo schema corporeo) sono i pazienti con **deafferentazione periferica**, che come vedremo in dettaglio nel capitolo 3 hanno subito un danno permanente alle fibre nervose che connettono i recettori periferici nella pelle e nei muscoli con il midollo spinale. La paziente G.L. studiata da Paillard [1999], ad esempio, era perfettamente in grado di indicare verbalmente quale parte del corpo era stata stimolata tattilmente da oggetti puntiformi o freddi, ed era inoltre in grado di indicare la posizione stimolata su un disegno schematico del corpo mostrando quindi una consapevolezza sia della presenza dell'evento tattile al corpo, sia della posizione di tale evento rispetto a una rappresentazione visiva e semantica del corpo. Al contrario, era ampiamente inaccurata nell'indicare manualmente il sito di stimolazione sul proprio corpo. Un esempio del secondo tipo di dissociazione (ovvero, danni a carico della sola immagine corporea) sono invece i pazienti con **deafferentazione centrale**, che a seguito di un danno corticale nelle regioni della corteccia somatosensoriale primaria e secondaria perdono la sensibilità somatosensoriale per uno o più distretti corporei controlaterali alla lesione. La paziente R.S. descritta da Paillard, Michel e Stelmach [1983], ad esempio, a seguito di una lesione nel lobo parietale sinistro non percepiva consapevolmente nessuna stimolazione statica a mano e avambraccio destri. Ciononostante, era sorprendentemente in grado di indicare correttamente la posizione della stimolazione tattile sulla mano destra con l'indice della mano sinistra. Nelle parole stesse della paziente emergeva tutta la sorpresa per questa dissociazione: «Io non capisco come questo possa accadere. Voi avete messo qualcosa [sulla mano destra], io non ho sentito nulla eppure sono andata esattamente lì con il mio dito [sinistro]» [Paillard 1999, 200].

La distinzione fra schema corporeo e immagine corporea, benché comunemente adottata da molti testi e articoli scientifici che discutono delle tematiche legate alla percezione e della rappresentazione del corpo, è però tutt'altro che unanime e condivisa. Alcuni autori hanno suggerito ad esempio che entrambe queste rappresentazioni del corpo debbano essere ricondotte a una rappresentazione unica e geneticamente determinata (la *neuromatrice* proposta da Mezack [1990]). Altri autori [Schwoebel e Coslett 2005; Sirigu *et al.* 1991] hanno invece proposto che sia possibile una suddivisione ulteriore all'interno del concetto di immagine corporea fra gli aspetti più strettamente visuopercettivi (**descrizione strutturale del corpo**) e quelli più strettamente legati alle conoscenze esprimibili su base verbale (**semantica del corpo**). Nuovamente questa ulteriore distinzione origina dall'osservazione di dissociazioni comportamentali fra popolazioni di pazienti neurologici: da un lato pazienti incapaci di pianificare ed eseguire azioni in assenza di disturbi motori elementari (*aprassia*; interpretata come esempio di schema corporeo deficitario), da un altro lato pazienti incapaci di localizzare parti del corpo sul proprio corpo, sul corpo dell'esaminatore o su un manichino (*autotopoagnosia*; interpretata come esempio di descrizione strutturale del cor-

po deficitaria) e da un altro ancora pazienti incapaci di denominare specifiche parti del corpo (*afasia* selettiva per parti del corpo; interpretata come esempio di un deficit a carico della conoscenza semantica del corpo). Per gli scopi di questo testo riteniamo sufficiente sottolineare nuovamente il concetto che le informazioni corporee presenti nella corteccia somatosensoriale primaria debbano necessariamente essere integrate con altre informazioni sensoriali per dare origine a rappresentazioni del corpo di più alto livello (per una trattazione dei molteplici frazionamenti della rappresentazione corporea rimandiamo il lettore alla lettura del contributo di de Vignemont [2007], in risposta all'articolo di Dijkerman e De Haan [2007]).

5. IL «MIO» CORPO: LA PERCEZIONE DI «OWNERSHIP»

Nella vita quotidiana la maggior parte di noi sa esattamente dove si collocano i **confini del corpo** e quali parti del corpo ci appartengano. Anche quando mescoliamo le nostre mani a quelle di altri (si pensi ad esempio alla situazione in cui due o più persone si stringono la mano) o quando viviamo in maniera fortemente empatica le sensazioni corporee provate da un nostro conoscente (si pensi ad esempio all'osservare una piccola operazione chirurgica come l'asportazione di una spina dalla mano di un amico) non cadiamo praticamente mai nell'incertezza di quali siano effettivamente le *nostre* parti del corpo. Tuttavia, come abbiamo visto nel paragrafo precedente e come vedremo ulteriormente nel capitolo 2 sui disturbi della rappresentazione corporea in pazienti cerebrolesi o con attacchi epilettici, questa certezza circa la proprietà di parti del corpo è tutt'altro che stabile. Il fatto che l'incertezza possa nascere anche dall'osservazione di una mano finta in una postura compatibile rispetto alla struttura del nostro corpo suggerisce che la spiegazione del fenomeno di *ownership* corporea vada cercata a un livello multisensoriale. Ma a quale delle molteplici rappresentazioni che abbiamo discusso va fatta risalire la sensazione proprietaria che ci consente di stabilire esattamente quale parte del corpo appartenga a noi e quale invece appartiene ad altri?

Ad oggi una risposta definitiva a questa domanda non c'è. Un'ipotesi plausibile però è che la sensazione di *ownership* derivi dalla localizzazione delle nostre sensazioni corporee all'interno di uno spazio definito [Jeannerod 2007; Martin 1995], e che questo spazio corrisponda allo schema corporeo, aggiornato in maniera dinamica istante per istante sulla base delle informazioni multisensoriali (principalmente visione, tatto e propriocezione) disponibili all'individuo [de Vignemont 2007].

Quanto alla prima parte di questa ipotesi, l'idea di fondo è che la regione di spazio che occupano le mie sensazioni corporee è unica, privata e non condivisibile. In altre parole è specifica per il mio Sé e preclusa al Sé di qualcun altro, e tutto ciò che cade all'interno di tale spazio deve quindi necessariamente appartenere al mio corpo. La seconda parte dell'ipotesi si basa invece su diversi argomenti. Anzitutto, lo schema corporeo è dedicato esclusivamente alla rappresentazione del proprio corpo, mentre l'immagine corporea è potenzialmente condivisa fra noi stessi e gli altri, in quanto può rappresentare tanto il nostro corpo quanto il corpo di chi ci sta attorno. Questa idea è illustrata bene dal fatto che persone con autotopoagnosia (tipicamente considerata un deficit dell'immagine corporea) falliscono nel localizzare parti del corpo sia sul proprio corpo sia sul corpo dell'esaminatore. In secondo luogo, lo schema corporeo rappresenta il corpo in azione: dunque il corpo come soggetto e come attore, piuttosto che come oggetto [*ibidem*]. In questo senso, nuovamente, favorisce fortemente l'assunzione di un punto di vista in prima persona rispetto al proprio corpo. Infine, un argomento empirico a supporto del fatto che la sensazione di *ownership* sia legata allo schema corporeo piuttosto che all'immagine corporea deriva dalle molte evidenze sperimentali che dimostrano un ruolo cruciale del controllo dell'azione nel generare un senso di *ownership* [van den Bos e Jeannerod 2002; Tsakiris, Prabhu e Haggard 2006; Pavani e Galfano 2007].

6. IL CORPO NELLA SUA INTEREZZA E IL SENSO DI SÉ

Dalla sensazione proprietaria del corpo deriva anche il nostro **senso del Sé**, l'esperienza conscia di essere un'entità olistica e distinta, capace di autocontrollo e controllo dell'attenzione, e che possiede – appunto – un corpo che occupa uno spazio e un tempo definiti [Blanke e Metzinger 2009]. Se da un lato il fenomeno dell'arto finto ci ha consentito di discutere adeguatamente concetti quali la formazione multisensoriale dello schema corporeo e il senso di *ownership*, esso appare ampiamente inadeguato per cogliere i meccanismi associati al senso del Sé. La nostra stessa introspezione ci dice, infatti, che il senso del Sé non è legato a singoli distretti corporei, ma dipende invece dal corpo nella sua globalità. I soggetti che esperiscono l'illusione dell'arto finto possono essere propensi a credere che una singola parte del corpo abbia cambiato posizione, ma non mettono in discussione il loro senso di sé. Analogamente, pazienti che a seguito di una lesione cerebrale negano l'appartenenza al loro corpo di un braccio (un fenomeno detto somatoparafrenia, che descriveremo più nel dettaglio

nel prossimo capitolo), non modificano di conseguenza anche la loro percezione del Sé. Un recente sviluppo della ricerca sui conflitti multisensoriali alla base dell'illusione dell'arto finto ha tuttavia aperto un nuovo filone di indagine che si prospetta molto più utile allo studio del senso del Sé.

Nel 2007, sulla rivista «Science», sono stati pubblicati due studi [Lenggenhager *et al.* 2007; Ehrsson 2007] che hanno indipendentemente mostrato che chiedendo a una persona di osservare in un visore per la realtà virtuale l'immagine del proprio corpo ripresa da una telecamera posta alcuni metri alle sue spalle, è possibile indurre nel soggetto la sensazione di osservare il proprio corpo in una posizione dello spazio diversa da quella fisicamente occupata. Questa idea sperimentale può essere fatta risalire alla fine del XIX secolo, quando G.M. Stratton [1899] osservò che il ripetuto utilizzo di occhiali prismatici in grado di proiettare di fronte a sé un'immagine del proprio corpo produceva bizzarre sensazioni di identificazione con quell'immagine percepita del corpo. Nelle situazioni illusorie studiate nel 2007, dove lo sdoppiamento avviene in un contesto immersivo di realtà virtuale, la schiena dei partecipanti all'esperimento veniva toccata diverse volte con un bastoncino, così che essi potevano sentire sulla loro schiena la stimolazione tattile, ma visivamente localizzare la stessa stimolazione di fronte a loro (creando un conflitto spaziale tra informazione tattile e visiva, in analogia a quanto succede con la mano di gomma).

Quando poi ai partecipanti veniva chiesto in quale punto della stanza si trovassero, quasi tutti hanno indicato la posizione virtuale. Questo suggerisce che il conflitto multisensoriale presente nella situazione sperimentale veniva risolto a favore della vista e a scapito di tatto e propriocezione. Inoltre, i partecipanti erano propensi a identificare il corpo visto come proprio in un questionario simile a quelli utilizzati per l'illusione dell'arto finto. In altre parole, questa illusione mostra che in alcune situazioni la visione può dominare le sensazioni corporee (tatto e propriocezione) ingannando il senso di appartenenza del nostro *intero corpo*, e causando una vera e propria sorta di **esperienza extracorporea** (*out-of-body experience*) come quelle di cui parliamo nel prossimo capitolo.

7. LE ESPERIENZE CORPOREE CONDIVISE E LA COGNIZIONE SOCIALE

Da tutto quanto abbiamo detto dovrebbe essere chiaro che la percezione multisensoriale del corpo svolge un ruolo fondamentale nei processi con cui rappresentiamo come siamo fatti e come agiamo nel mondo. Per questo motivo,

sembra lecito attendersi che l'esperienza multisensoriale del corpo possa essere importante anche per plasmare il modo in cui costruiamo la realtà sociale, e in particolare nella rappresentazione del nostro Sé nel contesto delle relazioni con i nostri simili: il nostro ruolo sociale, i gruppi di cui facciamo parte e così via. Nel rapportarci e nel comunicare con gli altri riceviamo costantemente informazioni multisensoriali sulla loro appartenenza a gruppi sociali, sulle preferenze e sui comportamenti legati alla cultura, sulle rappresentazioni e sulle credenze che riguardano la società nel suo complesso. Interpretare correttamente questi segnali è di fondamentale utilità per un corretto adattamento all'ambiente, ma potrebbe avere una funzione ancora più fondamentale nello sviluppo della cognizione umana. È stato proposto, infatti, che la rappresentazione di una **realtà condivisa** [Ecterhoff, Tory Higgins e Levine 2009] emerga nelle interazioni sociali grazie alla condivisione di esperienze percettive nel contesto della comunicazione collaborativa [Tomasello 2008]. Una rappresentazione di questo tipo potrebbe costituire l'elemento cruciale grazie al quale abbiamo accesso agli **stati mentali interni** dei nostri simili, costituendo una delle basi per la cognizione sociale. In una ricerca recente, ad esempio, Wiltermuth e Heath [2009] hanno mostrato che i partecipanti cui veniva richiesto di condividere esperienze percettive e motorie durante attività sincronizzate riportavano successivamente maggior attaccamento sociale nei confronti del loro gruppo e maggiore tendenza a cooperare in giochi economici, rispetto ai partecipanti che non svolgevano le attività sincronizzate.

La ricerca etologica, peraltro, ha da tempo riconosciuto l'importanza degli aspetti multisensoriali nella comunicazione sociale e alcuni studi hanno mostrato come il comportamento sociale di diverse specie animali sia il risultato dell'integrazione di molteplici segnali comunicativi [Partan e Marler 1999]. Per fare un esempio, le formiche *Aphaenogaster* emettono un particolare tipo di sostanza chimica (un feromone) quando vogliono richiedere la cooperazione di altre formiche per il trasporto di una preda. Tuttavia, se vogliono reclutare un maggior numero di collaboratori combinano l'emissione dei feromoni con una stridulazione (ovvero, un suono vibratorio prodotto dallo sfregamento di parti del corpo). Curiosamente, la stridulazione da sola non ha l'effetto di reclutare più formiche e dunque il comportamento sociale cooperativo sembra essere precisamente la conseguenza della combinazione dei due segnali sensoriali [Hölldobler 1999].

Nell'uomo lo studio dell'interazione fra gli aspetti multisensoriali della percezione e la cognizione sociale è solo agli inizi. Alcuni studi, ad esempio, si sono chiesti se la categorizzazione sociale può modulare l'interazione multi-

sensoriale. Uno studio rappresentativo di questa linea di ricerca è stato recentemente condotto da Andrea Serino e colleghi presso l'Università di Bologna [Serino, Giovagnoli e Làdavas 2009]. Gli autori di questa ricerca sono partiti da un precedente risultato [Serino, Pizzoferrato e Làdavas 2008] che mostrava come la capacità di un osservatore nel rilevare stimolazioni tattili al proprio volto è migliore quando egli guarda un video del proprio volto o il volto di un coetaneo dello stesso gruppo sociale che riceve analoghe stimolazioni tattili (rispetto a quando le stimolazioni tattili sono somministrate a un oggetto neutro, ad esempio la facciata di una casa). La domanda che si sono posti Serino e colleghi [2009] è stata se questa facilitazione multisensoriale della percezione tattile potesse essere mediata dalla categorizzazione sociale dell'altro, ovvero se potesse emergere in maniera selettiva quando il volto osservato apparteneva a un membro del proprio gruppo sociale (quello che in psicologia sociale è solitamente definito l'*ingroup*), rispetto a un membro di un gruppo sociale estraneo (l'*outgroup*). A questo scopo hanno confrontato gli effetti di facilitazione visuotattile quando l'appartenenza di gruppo era definita su base etnica, mostrando a soggetti caucasici o magrebini volti appartenenti alla medesima etnia o all'etnia diversa, e quando l'appartenenza di gruppo era definita su base politica, mostrando ad attivisti politici italiani democratici o conservatori i volti del leader della propria coalizione politica o della coalizione politica opposta (nello specifico i volti di Romano Prodi e di Silvio Berlusconi). I risultati hanno mostrato in maniera chiara che la facilitazione visuotattile emergeva in maniera selettiva quando il volto presentato veniva categorizzato come appartenente al proprio gruppo, fosse esso etnico o politico.

La ricerca appena descritta mostra quindi un'influenza della categorizzazione sociale su un fenomeno di percezione multisensoriale. Tuttavia, le ricerche etologiche a cui avevamo accennato in precedenza suggeriscono che potrebbe esistere anche una direzione opposta nella relazione fra queste due tematiche: ovvero, un effetto della percezione multisensoriale sul comportamento sociale. Uno studio condotto da Maria Paola Paladino e colleghi presso l'Università di Trento [Paladino *et al.* in stampa] ha di recente messo in evidenza precisamente questo tipo di relazione. Gli autori hanno mostrato che qualche minuto di stimolazione tattile al proprio volto, somministrata mentre il partecipante osserva un video che mostra un perfetto estraneo anch'egli toccato al volto, può modificare la categorizzazione sociale dell'estraneo. Se l'estraneo è toccato in maniera sincrona al partecipante, egli viene giudicato come più simile da un punto di vista fisico e da un punto di vista sociale. In particolare, viene categorizzato dall'osservatore come più vicino a sé dal punto di vista della relazione sociale

e più affine dal punto di vista dei tratti di personalità. Da notare che questi effetti di vicinanza sociale non emergono quando invece l'osservatore e l'estraneo vengono stimolati in maniera asincrona. In un'ottica multidisciplinare, ricerche come quella che abbiamo appena descritto hanno la potenzialità di offrire una nuova base empirica per l'osservazione antropologica che le azioni condivise e sincrone, quali ad esempio la danza, il canto o anche la marcia, possano giocare un importante ruolo di collante sociale [Fiske 2004]. Non a caso sono alla base di molti dei comportamenti rituali arcaici e moderni. La sfida per la ricerca, anche in questo ambito, è dunque quella di contestualizzare la cognizione sociale nell'ambiente multisensoriale nel quale viviamo ed esplorare le relazioni esistenti fra gli aspetti solo apparentemente concettuali del nostro comportamento e gli aspetti multisensoriali e motori che sono alla base della nostra interazione con il mondo.

CAPITOLO 2

Patologie della rappresentazione corporea

La letteratura neuropsicologica è ricca di descrizioni di patologie della rappresentazione del corpo. Queste patologie dimostrano che molti aspetti della percezione corporea possono essere compromessi in modo selettivo. La ricerca ha appena cominciato a chiarire quali siano i meccanismi funzionali e le basi neurali di queste anomale percezioni del corpo. A questo proposito, diversi elementi confermano l'importanza di affrontare lo studio di queste patologie e la riabilitazione adottando una prospettiva multisensoriale.

1. LE SOMATOPARAFRENIE

Percepire il nostro corpo è qualcosa che facciamo continuamente e che diamo per scontato. Ad esempio, troveremmo la cosa assai bizzarra se ci chiedessero se riteniamo che quella mano attaccata al nostro corpo sia veramente *nostra*. Tuttavia, la letteratura neuropsicologica riporta numerosi casi di pazienti cerebrolesi che mettono in evidenza come molto di quanto consideriamo ovvio delle nostre percezioni corporee non sia per nulla banale [Semenza 2001]. A seguito di una lesione cerebrale tipicamente nell'emisfero destro, ad esempio, alcuni pazienti possono negare che parti del loro corpo appartengano davvero a loro. Questo disturbo, battezzato **somatoparafrenia** da Gerstmann [1942] (per una rassegna vedi Vallar e Ronchi [2008]), colpisce le parti del corpo controlaterali alla lesione cerebrale, ovvero parti del corpo sul lato opposto rispetto all'emisfero cerebrale danneggiato. Frequentemente, la somatoparafrenia è associata alla **anosognosia** (incapacità di riconoscere i sintomi di una malattia che sono invece

ovvi) nei confronti di una paralisi di metà del corpo dovuta a un danno cerebrale al lobo parietale. Questi pazienti rifiutano dunque di riconoscere come propri gli arti sul lato lesionato del corpo. Tuttavia, sono stati descritti anche casi di somatoparafrenia per parti del corpo affette solo da parziale riduzione delle sensazioni tattili (**ipoestesia**).

Il fatto che le somatoparafrenie siano tipicamente a carico del lato sinistro del corpo potrebbe far supporre che l'integrità delle rappresentazioni corporee e la sensazione di identità e di appartenenza (*ownership*) del corpo siano generate dall'attività della metà destra del nostro cervello. Tuttavia questa conclusione, basata sulle rassegne dei casi riportati in letteratura, potrebbe essere messa in dubbio dal fatto che questo tipo di disturbo può essere osservato solo esaminando le risposte verbali dei pazienti. Questo potrebbe spiegare perché emerge molto meno in pazienti con lesioni cerebrali sinistre. Infatti, la comprensione e produzione linguistica sono spesso deficitarie in questi pazienti, mentre sono relativamente ben conservate in pazienti con lesioni all'emisfero destro. È possibile quindi che l'indagine verbale della somatoparafrenia vada a scontrarsi con le difficoltà linguistiche di questi pazienti.

I pazienti affetti da somatoparafrenia perdono la consapevolezza del possesso degli arti, di solito quelli superiori. I resoconti dei pazienti possono essere bizzarri: il braccio controlesionale appartiene a un altro paziente precedentemente trasportato in ambulanza, è stato dimenticato sul loro letto da un paziente precedente, qualcun altro è steso con loro nel letto, magari l'arto appartiene al medico che li sta curando. La cosa risulta ancor più sorprendente se si considera che questi pazienti possono anche non avere problemi psichiatrici, e spesso sono ampiamente coscienti di quanto siano bizzarre le loro percezioni. Nella maggior parte dei pazienti, inoltre, la somatoparafrenia è associata a **deficit di natura sensoriale**. In particolare, in questi pazienti è alterata la capacità di percepire e riconoscere la posizione del proprio corpo nello spazio senza il supporto della vista (ovvero la propriocezione). Mentre i disturbi di natura tattile o visiva non sono presenti in tutti i pazienti, la capacità di valutare la posizione di alcune parti del loro corpo è compromessa in tutti.

Ciononostante, i deficit di natura propriocettiva non sono sufficienti a spiegare il disturbo. Infatti, come abbiamo visto nel capitolo precedente, i segnali provenienti da altri canali sensoriali, quali la vista e il tatto, hanno anche un ruolo importante nella percezione di *ownership* del corpo (si pensi a quanto detto precedentemente in merito al fenomeno dell'arto finto), e potrebbero in linea di principio mitigare la sintomatologia associata alle somatoparafrenie. Sembra quindi plausibile ipotizzare che questi sintomi siano causati proprio

dalla mancata integrazione di sensazioni visive, tattili e propriocettive. A questo proposito è interessante notare quello che riportava una paziente studiata da Olsen [1937, citato da Nielsen 1938] quando le veniva mostrato che il braccio, di cui lei negava l'appartenenza, era effettivamente collegato con il resto del corpo: «[...] i miei occhi e le mie sensazioni non vanno d'accordo, e io devo credere alle mie sensazioni. Lo so che sembrano mie ma sento che non lo sono, e non posso credere ai miei occhi». Infine, il fatto che in questi pazienti vi sia un deficit di integrazione multisensoriale sembra essere confermato anche dalla sede delle lesioni che normalmente colpiscono questi pazienti e che riguardano, in particolare, aree cerebrali come la corteccia premotoria, lobo parietale posteriore e una serie di strutture sottocorticali come talamo, gangli della base e collicolo superiore coinvolte nell'integrazione multisensoriale.

2. L'ARTO FANTASMA

I pazienti affetti da somatoparafrenia hanno la sensazione che il proprio braccio o la propria gamba siano estranee al loro corpo. Altri pazienti, invece, possono avere la percezione di continuare a percepire un arto anche quando questo non c'è più. Il fenomeno è noto da molto tempo. Ad esempio, in uno degli ultimi capitoli dello straordinario *Moby Dick*, pubblicato per la prima volta nel 1851, Herman Melville immagina che il capitano Achab, amputato a una gamba, discorra con il falegname che gli sta preparando una nuova gamba di legno. Achab gli dice: «non farebbe troppo onore alla tua opera se, quando mi adatterò questa gamba che stai preparando, io mi sentissi ancora un'altra gamba perduta, quella di carne e d'ossa, insomma». A ciò il falegname replica: «Sì, ho sentito qualcosa di curioso a questo proposito, signore: come un uomo disalberato non perde mai interamente il senso del suo vecchio tronco, ma qualche volta lo sente che gli prude. Posso umilmente domandare se è davvero così, signore?». E Achab: «È così marinaio. Guarda, metti la tua gamba viva qui nel luogo dov'era la mia; così ora non c'è che una sola gamba visibile per l'occhio, pure per l'anima sono due. Dove tu senti formicolare la vita, lì, esattamente allo stesso punto, senza l'errore di un millimetro, sento anche io. È un indovinello?» (Herman Melville, *Moby Dick o la Balena*, Milano, Adelphi, 1987, pp. 492-493).

Melville dà una descrizione magistrale del fenomeno dell'**arto fantasma**, termine coniato da Mitchell, medico americano, che ne pubblicò una descrizione alla fine dell'Ottocento [Mitchell 1872]. Il fenomeno consiste in sensazioni

che provengono da un arto amputato e quindi inesistente. Il paziente dunque prova la sensazione che la parte del corpo amputata sia ancora presente. Infatti, ne avverte la posizione, spesso accusa sensazioni moleste e spesso dolorose, talora ne avverte addirittura il movimento, come se l'arto fosse ancora presente. Il fenomeno si riscontra nella quasi totalità dei soggetti amputati oppure sottoposti a denervazione totale dell'arto (per esempio, per sezione del plesso brachiale a livello spinale), indipendentemente dal livello intellettuale, dall'età del paziente e dalla sua consapevolezza della mutilazione subita. Le sensazioni fantasma non riguardano solo gli arti, ma possono essere originate, sia pure meno frequentemente, dall'asportazione di altre parti corporee, come i genitali maschili o i seni femminili. Dati ricavati da esperimenti di neuroimmagine sembrano suggerire che la rappresentazione corticale della parte amputata rimane attiva e che questa attività viene interpretata dal resto del cervello come proveniente dalla parte del corpo originariamente rappresentata in quella regione corticale.

Gli studi condotti negli anni '80 da Michael Merzenich [Merzenich *et al.* 1984] sui primati non umani e negli anni '90 da Vilayanur Ramachandran [Ramachandran, Rogers-Ramachandran e Cobb 1995] e da Salvatore Aglioti [Aglioti, Bonazzi e Cortese 1994] su persone amputate, sembrano suggerire che l'amputazione causi una ristrutturazione della mappa corticale del corpo che è presente nella corteccia somatosensoriale primaria (vedi fig. 1.1), a seguito di ciò i neuroni corrispondenti alla parte amputata sarebbero reclutati dalle aree corticali confinanti. Per esempio, toccando determinate regioni del volto, il paziente amputato riferisce la sensazione veridica localizzata nella parte toccata, ma anche una sensazione illusoria localizzata nell'arto fantasma, con una corrispondenza estremamente dettagliata. Secondo Ramachandran questa doppia sensazione, reale e illusoria, è dovuta alla **riorganizzazione delle aree corticali** normalmente attivate dalla stimolazione tattile della mano. A causa di tale riorganizzazione, nel momento in cui un arto viene a mancare, la corrispondente area corticale non riceve nessun input sensoriale e per questo è plausibile che gli input sensoriali provenienti dalla stimolazione tattile del viso oltre a raggiungere l'area del volto «invadano» anche l'area adiacente, che è appunto quella della mano (come si può notare guardando l'homunculus somatosensoriale di Penfield rappresentato in fig. 1.1). Questa ipotesi proposta tra gli altri da Ramachandran e Altschuler [vedi la loro rassegna pubblicata nel 2009] è stata testata da questi stessi autori utilizzando la magnetoencefalografia (MEG), una tecnica di neuroimmagine non invasiva che permette di studiare la funzionalità cerebrale attraverso la misurazione dell'attività elettromagnetica del cervello. Ramachandran e colleghi hanno mostrato come stimoli tattili presentati sul volto attivassero sia l'area del volto sia l'area della mano, suffragando così l'ipotesi di una riorganizzazione corticale a seguito dell'amputazione dell'arto.

Alcune testimonianze raccolte dallo stesso Ramachandran e riferite nel libro *La donna che morì dal ridere* [Ramachandran e Blakesee 1998] sottolineano bene questa relazione tra la topologia del corpo nella mappa somatosensoriale e il fenomeno dell'arto fantasma. Il racconto di Ramachandran riguarda le rappresentazioni di genitali e piedi, le cui relative aree cerebrali sono adiacenti nella corteccia somatosensoriale. Racconta il celebre neuropsicologo: «Il giorno dopo squillò il telefono. Era un ingegnere dell'Arkansas. – "Parlo con il dottor Ramachandran?" – "Sì" – "Ho letto delle sue ricerche su un quotidiano e ho trovato le notizie molto interessanti. Circa due mesi fa persi una gamba subito sotto il ginocchio, ma adesso mi sta succedendo qualcosa di strano. Vorrei il suo consiglio". – "Qual è il problema?" – "Ogni volta che ho rapporti sessuali, dottore, provo delle sensazioni al piede fantasma. Cosa dice? Il mio dottore pensa che è assurdo!" – "Non si preoccupi. Il fenomeno è forse dovuto al fatto che i genitali si trovano accanto ai piedi nell'area somatosensoriale della corteccia". – "Sarà anche così, dottore, ma vede lei forse non ha afferrato il concetto. Insomma io ora ho un orgasmo molto più potente di prima, perché non è limitato ai soli genitali"».

All'interno della cornice teorica del corpo come rappresentazione multisensoriale vale la pena citare anche l'osservazione secondo cui l'informazione visiva può contribuire a ridurre la consapevolezza dell'arto fantasma e a lenire il dolore che a seguito di tale percezione illusoria viene spesso percepito. Un esempio di riduzione nella sensazione illusoria prodotta dall'arto amputato è il fenomeno dell'**evitamento dell'ostacolo** (*obstacle shunning*). In questo caso, la sensazione dell'arto fantasma viene attenuata quando un oggetto arriva a occupare lo spazio che sembra essere occupato dall'arto fantasma. In generale, un approccio multisensoriale alla costruzione della rappresentazione corporea può offrire spunti importanti per la pratica riabilitativa in presenza di alterazioni della rappresentazione corporea. Secondo alcune osservazioni di Ramachandran, ad esempio, in pazienti amputati la semplice richiesta di osservare in uno specchio l'immagine riflessa dell'arto rimanente (che appare nella posizione in cui avrebbe dovuto trovarsi l'arto amputato) può dare al paziente amputato l'impressione di poter controllare l'arto fantasma [vedi Ramachandran e Altschuler 2009, per una rassegna]. Ramachandran ha battezzato questo approccio terapeutico **terapia dello specchio** (*mirror therapy*).

Ci sono stati alcuni tentativi di utilizzare la terapia dello specchio per fini riabilitativi. Come detto in precedenza, il fenomeno dell'arto fantasma è vissuto

come un'esperienza dolorosa da una percentuale molto alta di persone che subiscono amputazioni agli arti (50-80% secondo stime fatte da Jensen e Nikolajsen [1999]). Per esempio, chi ha perso il proprio braccio può, talvolta, percepire l'arto fantasma come paralizzato, la mano chiusa a pugno, impossibile da aprire, e questa sensazione provoca dei dolorosissimi crampi. La sensazione dolorosa associata all'arto fantasma è stata definita **dolore fantasma**. Il dolore fantasma può durare per anni, può essere continuo o intermittente, e diversi approcci sono stati adottati per curarlo, o perlomeno lenirlo: dall'ipnosi a interventi di chirurgia cerebrale. Tuttavia, queste terapie si sono mostrate inefficaci o molto poco efficaci. Dopo tutto, come è possibile trattare il dolore a un arto che non c'è più?

Ramachandran e i suoi colleghi sostengono che è possibile dare sollievo alle persone amputate utilizzando la cosiddetta *mirror box* (letteralmente scatola-specchio), una scatola divisa in due parti, una a destra e una a sinistra della linea mediana del corpo del paziente che vi si siede di fronte. Il paziente inserisce frontalmente nella *mirror box* l'arto sano e l'arto amputato, in modo da posizionarli a destra e a sinistra dello specchio. Guardando l'arto sano riflesso nello specchio il paziente ha l'impressione di vedere di nuovo il proprio arto amputato. Ramachandran e colleghi hanno descritto alcuni pazienti che riportavano una riduzione del dolore fantasma dopo avere effettuato movimenti simmetrici della mano reale (che, vista nello specchio, riproduce l'informazione visiva dell'arto mancante) e dell'arto fantasma (di fatto inesistente). Guardando la mano riflessa nello specchio, il paziente riportava la percezione illusoria che l'arto fantasma stesse effettivamente muovendosi come desiderato. Secondo Ramachandran, sperimentando questo fenomeno di cattura (o dominanza visiva) i pazienti riferiscono una diminuzione del dolore fantasma. La dominanza visiva responsabile delle illusioni multisensoriali descritte precedentemente potrebbe quindi essere sfruttata per modulare sensazioni corporee altrimenti anomale.

3. L'ARTO FANTASMA SOVRANNUMERARIO

Percezioni corporee fantasma possono presentarsi, a seguito di lesioni cerebrali, anche quando non vi è parziale o totale deafferentazione di segmenti corporei. Esistono in letteratura numerosi casi di pazienti cerebrolesi che riferiscono di percepire la presenza di più arti rispetto a quelli reali. Queste percezioni vengono definite dell'**arto fantasma sovrannumerario** (o degli arti fantasma sovrannumerari a seconda che il paziente percepisca uno o più arti oltre ai quattro normalmente percepiti ed effettivamente presenti). In alcuni casi, i pazienti dichiarano di possedere anche quattro braccia e sei gambe. In uno dei più dettagliati resoconti del fenomeno, Ehrenwald [1930] riportava il caso di un paziente, con emiplegia all'arto superiore e inferiore, che si lamentava di avere un «nido di mani nel suo letto» e chiedeva insistentemente ai medici che lo assistevano di amputare quelle mani.

Le lesioni che causano il fenomeno possono essere localizzate in aree diverse del cervello. Per esempio, lesioni parietali, soprattutto all'emisfero cerebrale di destra, sembrano produrre il fenomeno a seguito dell'alterazione nell'afferenza somatosensoriale e nella consapevolezza corporea. Invece, cause di natura motoria sarebbero quelle che originano il fenomeno a seguito di lesioni alle aree temporomesiali del cervello. In questo caso, l'attivazione della regione premotoria (deputata alla programmazione dei movimenti) continuerebbe anche dopo l'esecuzione del movimento, creando la sensazione illusoria che vi sia un arto fantasma che si sta per muovere, mentre il movimento è già stato fatto dall'arto reale. Un'altra differenza che distingue le sensazioni fantasma conseguenza di lesioni cerebrali dalle sensazioni fantasma procurate da amputazioni degli arti è il diverso grado di realtà che viene attribuito ai due fenomeni dalle persone che li provano. Le percezioni degli arti fantasma sono molto più vivide e reali nelle persone amputate rispetto alle persone cerebrolese. Tuttavia, è interessante notare che la percezione dell'arto fantasma sovrannumerario può essere più o meno vivida a seconda di quanto sia alterata la percezione dell'arto controlaterale alla lesione cerebrale.

Infine, l'esperienza dell'arto fantasma sovrannumerario si distingue da quella descritta da persone amputate per il suo contenuto percettivo multisensoriale. In particolare, nessun amputato ha mai riferito di aver visto l'arto fantasma, mentre esistono pazienti cerebrolesi che hanno riferito di aver avuto esperienze sia visive sia posturali-cinestesiche del loro arto sovrannumerario. Per esempio, un paziente descritto da Miyazawa *et al.* [2004] diceva di vedere chiaramente le sue braccia e gambe sovrannumerarie, anche se lo stesso non avveniva per le mani e piedi corrispondenti. Tuttavia, si ritiene che questi fenomeni riflettano influenze cognitive sui processi percettivi, più che vere e proprie alterazioni a carico dei processi di integrazione multisensoriale [Brugger 2006].

4. I FENOMENI AUTOSCOPICI

Le alterazioni della rappresentazione corporea non si limitano solamente a singoli arti, o parti del corpo, ma possono riguardare l'intero corpo. Ciò accade

nel caso dei **fenomeni autoscopici**. Si tratta di percezioni illusorie, caratterizzate dall'impressione di percepire il proprio corpo occupare simultaneamente due posizioni nello spazio: la posizione reale e una posizione virtuale. I fenomeni autoscopici comprendono le esperienze extracorporee (*out-of-body experience*), l'allucinazione autoscopica e infine l'*heautoscopy* (letteralmente «visione di se stessi»). Queste tre varianti dei fenomeni autoscopici differiscono per le caratteristiche fenomeniche della percezione corporea illusoria coinvolta (per una rassegna, vedi Brugger [*ibidem*]).

Nel caso delle **esperienze extracorporee**, i pazienti riferiscono di avere avuto esperienze extracorporee quando si trovano in posizione supina e di avere avuto la sensazione di essere fuori dal loro corpo. In particolare, di sentirsi sospesi nell'aria in posizione parallela e speculare a quella del loro corpo, in realtà sdraiato. Come se fluttuassero nello spazio e, da questa prospettiva aerea, potessero vedere sia il proprio corpo, sia le persone e gli oggetti che li circondano. Nel caso dell'**allucinazione autoscopica**, i pazienti riferiscono di vedere un duplicato di se stessi, ma in questo caso senza avere l'impressione di uscire dal proprio corpo. Infatti, questi pazienti raccontano di vedere una copia del proprio corpo di fronte a loro, in posizione speculare. Quindi, nel caso dell'esperienza extracorporea, i pazienti vedono il loro corpo reale dalla prospettiva del corpo virtuale, mentre nel caso dell'allucinazione autoscopica è il corpo virtuale a essere visto dalla prospettiva del corpo reale.

Un altro elemento che differenzia esperienze extracorporee e allucinazioni autoscopiche è che, al contrario dell'esperienza extracorporea, l'allucinazione autoscopica viene di solito percepita quando si è in posizione eretta o si sta seduti. Lo stesso accade nel caso della cosiddetta **«heautoscopy»**. Quest'ultima variante ha però manifestazioni che la rendono un fenomeno intermedio tra l'esperienza extracorporea e l'allucinazione autoscopica. Infatti, i pazienti che hanno questo tipo di illusione riferiscono di poter vedere una copia di se stessi proprio davanti a sé, ma è difficile per loro stabilire se si sentono presenti esclusivamente nel proprio corpo reale (come nell'allucinazione autoscopica) oppure se si sentono trasferiti nella copia di loro stessi (come nell'esperienza extracorporea). In sostanza è come se si sentissero simultaneamente in entrambe le posizioni: quella reale e quella virtuale. Il fenomeno della *heautoscopy* si differenzia dagli altri due anche per un altro aspetto, ovvero la spiacevolezza dell'esperienza vissuta dal paziente. Contrariamente a quanto accade nell'esperienza extracorporea e nell'allucinazione autoscopica, nella *heautoscopy* il paziente spesso riferisce una vera e propria situazione conflittuale fra sé e la propria immagine.

I fenomeni autoscopici appena descritti hanno come caratteristica comune quella di venire raccontati dai pazienti come esperienze che si manifestano visivamente. Queste illusioni sono caratterizzare dalla percezione di una doppia copia del proprio corpo: una che osserva l'altra (anche se la distinzione tra parte attiva e passiva è più sfumata nell'*heautoscopy*). Questo ha portato a ipotizzare che le cause di questi fenomeni siano da ricercare prevalentemente in un'alterazione di natura visiva. Tuttavia, i fenomeni autoscopici vengono vissuti dai pazienti come fortemente reali, diversi da ciò che si prova durante per esempio i sogni. Ciò fa supporre che non si tratti puramente di un'esperienza visiva. Queste considerazioni hanno suggerito che anche nel caso dei fenomeni autoscopici vi potrebbero essere delle anomalie nell'integrazione delle informazioni sensoriali coinvolte nella rappresentazione del corpo, in particolare, della sua forma e della posizione che occupa nello spazio. A favore di questa ipotesi, ci sono anche i dati relativi alle sedi delle lesioni cerebrali che colpiscono i pazienti che sperimentano queste illusioni di duplicazione corporea. Nella gran parte dei casi, le aree cerebrali danneggiate si collocano nella **giunzione temporo-parieto-occipitale**, una zona del cervello che ha un ruolo molto importante nell'integrazione multisensoriale.

Vari autori (per una rassegna vedi Blanke e Mohr [2005]) hanno proposto che vi sia un'alterazione della percezione corporea che riguarda l'integrazione di informazioni propriocettive-cinestesiche e vestibolari. Un coinvolgimento della percezione della posizione del proprio corpo e del suo movimento (propriocezione-cinestesia) sembra essere suggerito dal fatto che molti pazienti riportano di percepire una concomitanza di movimenti tra il loro corpo reale e il duplicato. A questo si aggiunga che all'interno dei fenomeni autoscopici la copia del corpo sembra assumere posizioni differenti a seconda dell'illusione percepita: posizione supina nell'esperienza extracorporea e invece posizione eretta nell'allucinazione autoscopica e nell'*heautoscopy*. Infine sembra essere coinvolto nella generazione dei fenomeni autoscopici anche il senso vestibolare che fornisce informazioni sull'orientamento del corpo nello spazio tridimensionale. Molti di questi pazienti mostrano disturbi di natura vestibolare come la percezione di sentire il proprio corpo fluttuare o volare. Queste sensazioni sono soprattutto presenti nell'esperienza extracorporea, in proporzione minore nell'*heautoscopy*, e raramente presenti nelle allucinazioni autoscopiche.

5. LA SINDROME DELLA MANO ANARCHICA

Ulteriori dati a favore della stretta relazione fra pianificazione dei movimenti e rappresentazione del corpo provengono dal curioso disturbo noto come **sindrome della mano anarchica.** Le persone che ne sono affette sono vittime di un conflitto tra il controllo sull'azione di una delle loro mani e quello che la mano effettivamente fa. Insomma, una delle mani è fuori dal controllo della volontà del paziente come se questa mano stessa avesse una mente e una volontà propria. Chi ama il cinema di Stanley Kubrick non avrà mancato di riconoscere il disturbo che, nella finzione cinematografica, affliggeva il folle scienziato impersonato da Peter Sellers nel film *Il dottor Stranamore*. In una scena esilarante nella parte finale del film, infatti, il braccio destro del dottore inizia autonomamente a fare il saluto nazista mentre questi illustra allo stato maggiore un piano per la sopravvivenza della razza umana dopo un conflitto nucleare. Il braccio scatta a più riprese verso l'alto mentre Stranamore-Sellers si contorce cercando di fermarlo e riportarlo giù con la mano sinistra.

Come nel caso del dotto Stranamore, i pazienti affetti da sindrome della mano anarchica sono perfettamente consapevoli dell'anomalia, ma non sono in grado di inibirne la manifestazione. Ecco il resoconto fatto da Sergio Della Sala [Della Sala, Marchetti e Spinnier 1991], uno dei maggiori studiosi del fenomeno: «Una sera invitammo la nostra paziente, la signora G.P., a cena con la sua famiglia. Stavamo discutendo le implicazioni della sua condizione di salute per lei e i suoi familiari, quando improvvisamente e con grande costernazione da parte sua, la mano sinistra di G.P. prese alcune lische di pesce avanzate e le portò alla bocca. Poco dopo, afferrò il gelato che suo fratello stava leccando. La mano destra intervenne immediatamente a mettere le cose a posto e come risultato della lotta il gelato cadde sul pavimento». Parking [1996] riporta il caso di un paziente che aveva problemi a scegliere i canali televisivi perché «non appena la mano destra sceglieva un canale, la mano sinistra voleva premere un altro bottone».

I pazienti affetti da mano anarchica sono dunque vittima di questi episodi in cui la mano inizia da sé a svolgere azioni non volute e spesso anche imbarazzanti. Va notato tuttavia che costoro non riferiscono che la mano in questione non sia la «loro» mano. Presumibilmente, dunque, il sintomo non va collegato a un disturbo nei processi responsabili della percezione di *ownership* del corpo (vedi cap. 1), ma quelli coinvolti nell'elaborazione del movimento intenzionale. In particolare, la sindrome della mano anarchica sembra associata a lesioni all'**area supplementare motoria** (SMA, *Supplementary Motor Area*) [Della Sala, Marchetti e Spinnler 1994]. Numerosi dati neurofisiologici e di neuroimmagine indicano che SMA è coinvolta nella conversione di intenzioni in azioni e nella selezione, e va tenuta distinta dalla **corteccia premotoria** (PMC, *Pre-Motor Cortex*), responsabile invece dei movimenti eseguiti in risposta a stimoli esterni [Blakemore, Wolpert e Frith 2002; Mushiake, Masahiko e Tanji 1991]. La normale interazione con l'ambiente richiede, evidentemente, un corretto equilibrio fra azioni in risposta a stimoli esterni e azioni selezionate a partire da decisioni interne. A seguito di una lesione a carico di SMA, questo equilibrio verrebbe alterato, cosicché la pianificazione e l'inizio del movimento della mano controlaterale alla lesione possono essere provocati autonomamente da stimoli esterni elaborati dalla PMC. Ciò suggerisce dunque che l'impressione di «avere deciso» di svolgere un'azione (*agency*) potrebbe essere in realtà un processo di inibizione delle azioni alternative in quel momento attivate, un'idea su cui torneremo nella parte finale del prossimo capitolo.

L'ipotesi è confermata dall'esame di pazienti con lesioni bilaterali alle SMA, uno dei quali è stato descritto proprio dallo stesso gruppo di Della Sala [Boccardi *et al.* 2002]. Questo paziente manifestava una necessità impellente di utilizzare gli oggetti vicini alle mani, un sintomo conosciuto come **comportamento di utilizzazione.** Per esempio, durante un test, avendo scorto una mela e un coltello sul tavolo, il paziente la sbucciò e se la mangiò. L'esaminatore, curioso per il comportamento tenuto dal paziente, gli chiese perché avesse mangiato la mela. Il paziente rispose semplicemente: «Perché era qui». In sostanza, il paziente affetto da comportamento di utilizzazione risulta incapace di inibire qualsiasi tipo di comportamento suggerito dall'ambiente. In questo caso, essendo presente una lesione bilaterale, il sintomo interessa entrambi gli arti e il paziente ha scarsa consapevolezza del fatto che il suo comportamento è socialmente inappropriato. Nel caso della mano anarchica, invece, la lesione unilaterale scatena un conflitto tra le due mani: quella controlaterale alla lesione sfugge al controllo della volontà del paziente, mentre quella ipsilaterale rimane sotto il suo controllo. Proprio a causa di questo conflitto, il paziente rimane consapevole dell'inappropriatezza delle azioni svolte dalla sua mano anarchica. L'area SMA non lesionata continua infatti a funzionare normalmente, e in questa maniera presumibilmente il paziente continua ad avere un minimo di accesso al proprio sistema di controllo.

CAPITOLO 3

Muovere il corpo

Oltre a contribuire alla percezione del corpo statico, i processi multisensoriali sono coinvolti nella guida del corpo in movimento. Questo è evidente dall'esame di compiti motori semplici, quali lo spostamento della mano verso un oggetto, l'azione di raggiungimento-prensione, la deambulazione e il mantenimento della postura. Inoltre preziose informazioni sulla percezione e la guida dei movimenti del corpo derivano dallo studio dei pazienti deafferentati delle sensazioni somatiche.

1. LE TRASFORMAZIONI SENSOMOTORIE

La rappresentazione percettiva del corpo non serve solo a sapere come siamo fatti, ma anche a guidare i nostri movimenti. Per gestire il corpo in movimento, infatti, la mente deve effettuare operazioni che vengono chiamate **trasformazioni sensomotorie**. Il termine fa riferimento all'insieme di operazioni mentali necessarie a trasformare informazioni sensoriali in azioni, ossia in sequenze di movimenti organizzate allo scopo di raggiungere determinati scopi. Si tratta di operazioni mentali complesse, che coinvolgono la codifica dell'informazione sensoriale in ingresso, la sua interpretazione al fine di evidenziare gli aspetti rilevanti per l'azione dell'effettore (la parte del corpo utilizzata per svolgere il movimento) e dell'oggetto bersaglio, la loro rappresentazione in sistemi di coordinate adatti alla rappresentazione del movimento, la specificazione dei comandi motori necessari a iniziare il movimento, l'eventuale monitoraggio della nuova informazione disponibile durante lo svolgimento dell'azione e la modifica dei comandi motori allo scopo di correggere eventuali errori rilevati. Nella

fase preparatoria dell'azione, inoltre, una trasformazione sensomotoria può essere influenzata da conoscenze dichiarative o procedurali provenienti da sistemi di memoria a lungo termine. Queste contribuiscono alla scelta della tipologia di azione da compiere (ad esempio, afferrare un oggetto con una mano o con due mani; oppure afferrarlo dal lato lungo invece che dal lato corto) e, secondo alcuni ricercatori, in alcune condizioni possono influenzare anche la definizione dei parametri dei movimenti elementari che costituiscono l'azione stessa (ad esempio, quanto aprire le dita per andare a prendere un oggetto). Tenuto conto della complessità di questa sequenza di operazioni e del gran numero di sistemi mentali coinvolti, non è sorprendente che la guida percettiva del movimento sia uno dei processi mentali in cui una prospettiva multisensoriale si rivela maggiormente fruttuosa.

In questo capitolo ci occuperemo dei processi multisensoriali coinvolti nella guida del movimento in quattro situazioni: lo spostamento della mano verso un oggetto bersaglio, il movimento di raggiungimento-prensione, la deambulazione e il mantenimento della postura. Si tratta, come si vede, di azioni piuttosto elementari che si prestano bene a essere studiate in esperimenti ben controllati con apparecchiature in grado di registrare in modo preciso le durate e le traiettorie caratteristiche dei movimenti di risposta. Ciò non significa naturalmente che la ricerca non si sia interessata anche di situazioni ben più complesse.

Ad esempio, in una recente rassegna Gray [2008] ha discusso una serie di risultati empirici che testimoniano del ruolo dell'informazione multisensoriale nell'attività sportiva e nel pilotaggio di aerei e di automobili. Uno di questi riguarda il controllo della direzione di volo di un aereo in condizioni in cui venga improvvisamente aumentata l'accelerazione in avanti del velivolo. Come viene illustrato nella figura 3.1, in queste condizioni la percezione del movimento del corpo del pilota dipende, oltre che dall'informazione visiva, da due tipi di stimolazione meccanica: la forza di gravità (g), risultante, a livello muscolare e tattile, nella sensazione di pressione sul sedile, e la forza inerziale (I) dovuta all'accelerazione in avanti, registrabile dal sistema vestibolare e, in qualche misura, di nuovo dalla pressione sul sedile. Si noti che la risultante di queste due forze (R) è in una direzione coerente con un movimento verso l'alto. Immaginiamo ora che l'informazione visiva sulla direzione di volo sia degradata o assente, come nel volo notturno o in presenza di nuvole. Come è facile immaginare, i piloti possono essere tratti in inganno percependo movimento verso l'alto (la cosiddetta **illusione «heads up»**), e quindi mettere in atto una risposta compensativa, portando la cloche in avanti e spingendo l'aereo in picchiata, con conseguenze potenzialmente pericolose. In questa sede non potremo occuparci,

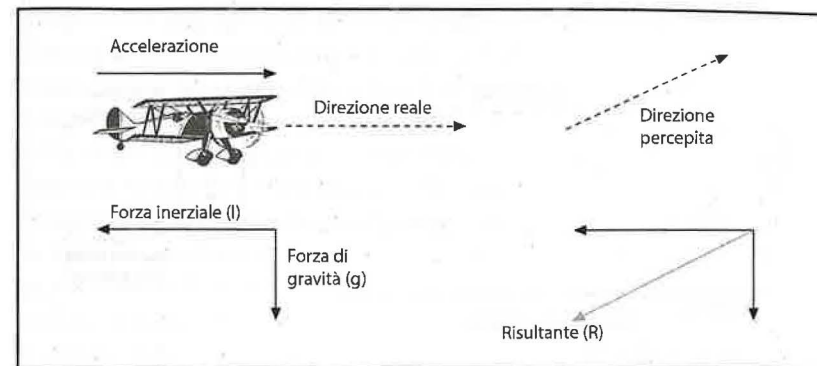



fig. 3.1. L'illusione *heads up* nel pilotaggio di un aereo. In condizioni di aumento dell'accelerazione in avanti del velivolo, i sistemi vestibolare e tattile registrano la risultante della forza inerziale e della forza di gravità. In assenza di ulteriori informazioni sulla direzione di volo si può produrre la percezione erronea che l'aereo stia volando verso l'alto.

fonte: Gray [2008].

sia per ragioni di spazio sia per chiarezza di esposizione, di tutte le situazioni naturalistiche in cui i processi multisensoriali di guida del movimento entrano in gioco. Le situazioni più semplici che verranno analizzate ci consentiranno tuttavia di illustrare alcuni principi che riteniamo possano essere generalizzati anche a contesti più naturali o complessi.

2. MUOVERE LA MANO

Il semplice spostamento della mano verso una posizione che funge da bersaglio costituisce un punto di partenza particolarmente adatto per studiare i processi percettivi e motori coinvolti nella guida delle azioni. Si tratta di un'azione semplice (coinvolge, tipicamente, solo le articolazioni della spalla e del gomito) e sulla quale non è difficile ottenere dati interessanti anche senza disporre di tecnologie sofisticate. Infatti, già alla fine del XIX secolo Woodworth [1899], nella sua tesi di dottorato alla Columbia University, riuscì a dimostrare alcuni dei principi fondamentali che descrivono il funzionamento dei **sistemi di guida dei movimenti**.

Per raccogliere dati sulla precisione dello spostamento della mano, Woodworth costruì un apparato che faceva scorrere una striscia di carta a velocità costante su un rullo. I partecipanti all'esperimento dovevano tenere in mano

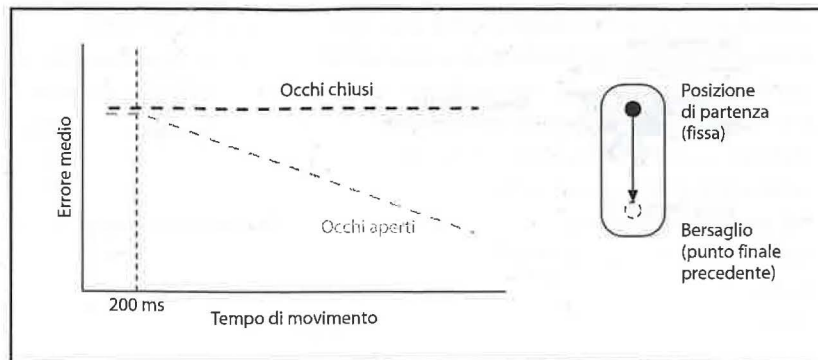



fig. 3.2. Errore medio nello spostamento della mano con o senza informazione visiva durante il movimento. A ogni prova, i partecipanti spostavano una penna su un foglio di carta che scorreva a velocità costante su un rullo. Il bersaglio era costituito dalla posizione finale, memorizzata, del movimento precedente.

fonte: Adattata da Woodworth [1899].

una penna ed effettuare ripetuti movimenti laterali su questa striscia di carta, partendo da un punto fisso e con la consegna che ognuno di questi spostamenti avrebbe dovuto avere la stessa lunghezza dello spostamento precedente. I partecipanti compivano questi movimenti a velocità variabili, sincronizzandosi con lo scandire di un metronomo. Inoltre Woodworth chiese di svolgere questo compito in due condizioni diverse, ossia guardando la propria mano durante lo spostamento o chiudendo gli occhi prima di iniziare lo spostamento. Analizzando i tracciati ottenuti in queste due maniere, Woodworth si rese conto che l'errore medio nel riprodurre l'ampiezza precedente variava in funzione del tempo impiegato per lo spostamento. Per tempi di movimento molto brevi (al di sotto dei 150 millisecondi) l'errore era piuttosto grande ma era equivalente nelle due condizioni. Per tempi di movimento più grandi, invece, l'errore tendeva a essere sempre più piccolo quanto più lento era il movimento, ma solo nella condizione in cui i partecipanti potevano vedere la loro mano. Nella condizione in cui il movimento veniva completato a occhi chiusi, invece, l'errore rimaneva sostanzialmente identico sia con movimenti veloci sia con movimenti lenti (fig. 3.2).

Da questi risultati Woodworth concluse che il sistema che guida lo spostamento della mano è scomponibile in due fasi. Nella prima fase, che chiamava **impulso iniziale** (*initial impulse*), il sistema invia un impulso o comando motorio programmato prima di iniziare il movimento. Nella seconda fase, che chiamava **monitoraggio** (*current control*), il sistema utilizza il segnale visivo afferente per correggere la traiettoria del braccio durante il movimento, operando

gli opportuni aggiustamenti allo scopo di minimizzare l'errore nel raggiungere il bersaglio. Questi aggiustamenti saranno ovviamente tanto più efficaci quanto maggiore è il tempo disponibile prima della fine del movimento. Quando invece il segnale visivo afferente non è disponibile, perché il movimento viene fatto a occhi chiusi, questi aggiustamenti non sono possibili e quindi l'errore finale riflette la posizione programmata prima di iniziare il movimento. Infine, quando il tempo di movimento è inferiore al tempo necessario per completare almeno un ciclo di correzione, tenere gli occhi aperti non comporta alcun beneficio e quindi l'errore nelle due condizioni rimane lo stesso. Da quest'ultima osservazione Woodworth concluse dunque che il tempo necessario per elaborare il segnale visivo afferente e correggere il movimento si aggira intorno ai 150 millisecondi.

I risultati di Woodworth sono considerati ancora oggi una pietra miliare nello studio dei processi percettivi e motori coinvolti nella guida del movimento. Ad esempio, avremo modo di citare nuovamente la tecnica che prevede l'esclusione del segnale visivo durante lo svolgimento dell'azione nel discutere molti altri risultati sperimentali descritti in questo libro. Inoltre le due fasi identificate da Woodworth corrispondono alla fondamentale distinzione, ancora oggi utilizzata, fra **guida a circuito aperto**, ossia basata sul **programma motorio** formulato prima di iniziare il movimento, o **guida a circuito chiuso**, successiva alla «chiusura» del ciclo di elaborazione del segnale visivo e di correzione del movimento (per un approfondimento su questo e altri concetti utilizzati nello studio della psicologia del controllo motorio rimandiamo alla monografia di Nicoletti e Borghi [2007]).

A dispetto di alcune debolezze metodologiche – ad esempio, la posizione bersaglio, memorizzata, è più ambigua di un bersaglio effettivamente presentato al partecipante, e l'esclusione del segnale visivo oggi viene controllata in maniera più precisa utilizzando particolari occhiali controllati via software – i risultati di Woodworth sono stati sostanzialmente replicati da studi più recenti [Keele e Posner 1968], persino nella sua stima del tempo necessario per completare il primo ciclo di aggiustamento. Adottando una prospettiva multisensoriale gli esperimenti di Woodworth hanno tuttavia un'altra limitazione. Per la natura stessa del compito utilizzato, in cui il bersaglio da raggiungere è specificato dalla visione della posizione finale della mano nella prova precedente, o dalla memoria recente di quella posizione (nella condizione a occhi chiusi), il compito di Woodworth non consente di valutare il contributo di informazioni potenzialmente disponibili in altri canali sensoriali per la guida del movimento. In particolare, sembra plausibile aspettarsi che anche le informazioni tattili e

propriocettive sulla posizione della mano possano essere utilizzate per dirigere il movimento, specialmente quando l'afferenza visiva venga esclusa. Il compito di Woodworth non si presta bene a valutare questo contributo, perché la posizione del bersaglio è in qualche maniera ambigua e dipende dal ricordo del partecipante.

Molte ricerche più recenti hanno invece dimostrato in maniera molto convincente che, in condizioni diverse, i segnali extravisivi sulla posizione della mano o, in alcuni esperimenti, dello stesso bersaglio, contribuiscono in maniera determinante a modulare la traiettoria del braccio e la posizione di arrivo della mano. Questi esperimenti suggeriscono inoltre che la rappresentazione corporea utilizzata da questi processi multisensoriali non è riconducibile a una mera sequenza di comandi motori, ma dipende proprio da una rappresentazione più astratta e flessibile del corpo in movimento. I metodi utilizzati sfruttano, anche se in maniere diverse, un paradigma sperimentale che viene chiamato **adattamento prismatico**. Il fenomeno dell'adattamento prismatico, noto da moltissimo tempo, si riferisce all'utilizzo di occhiali su cui sono montati prismi in grado di deviare i raggi luminosi, producendo distorsioni sistematiche nella struttura spaziale della luce che raggiunge la retina. Ad esempio, sono state studiate situazioni in cui i prismi ruotano l'immagine retinica di 180 gradi, invertendo la direzione alto-basso [Stratton 1897a; 1897b] oltre a distorsioni meno drastiche come quelle prodotte da prismi che spostano i raggi lateralmente di alcuni gradi o incurvano la proiezione di un contorno rettilineo [Kohler 1962]. L'aspetto interessante di queste esperienze è che, mentre all'inizio la distorsione causa errori macroscopici nel direzionamento dei movimenti, con l'abitudine questi errori scompaiono, suggerendo che il sistema è in grado di attuare una **ricalibrazione** delle trasformazioni sensomotorie coinvolte. Ad esempio, se il prisma sposta la proiezione di un oggetto di cinque gradi di angolo visivo verso sinistra, al primo tentativo di raggiungimento la mano percorre una traiettoria che devia marcatamente in quella direzione. Diverse ricerche hanno mostrato che questo errore iniziale corrisponde bene alla media delle direzioni specificate dalla visione, distorta, e dal segnale propriocettivo, veridico [Rossetti, Desmurget e Prablanc 1995]. In seguito all'adattamento, tuttavia, questo errore iniziale scompare e la mano raggiunge il bersaglio nonostante la distorsione.

Proprio utilizzando l'adattamento prismatico, già alcuni decenni fa Harris [1963] ha riportato un risultato particolarmente rilevante per valutare l'influenza dei diversi canali sensoriali nella guida dello spostamento della mano. L'elegante esperimento di Harris sfruttava un'altra caratteristica dei fenomeni di adattamento, il fatto che, una volta completata la ricalibrazione che consente al

TAB. 3.1. *Aftereffect* (differenza media in cm fra la prestazione prima e dopo l'adattamento prismatico) nelle sei condizioni studiate da Harris [1963]

	MANO ADATTATA	MANO NON ADATTATA
Bersaglio visivo	5,8 cm	1 cm
Bersaglio acustico	5,1 cm	-0,5 cm
Movimento in avanti	5,6 cm	0,5 cm

sistema di compensare le distorsioni, se la fonte di distorsione viene rimossa (il partecipante si toglie i prismi, ad esempio) nel periodo immediatamente successivo si osservano **effetti postumi** (i cosiddetti *aftereffect*) nella direzione opposta a quella dell'adattamento. Detto altrimenti, se il sistema ha «imparato» che per raggiungere un bersaglio spostato otticamente di cinque gradi verso sinistra occorre guidare il movimento nella direzione cinque gradi verso destra, tolti i prismi inizialmente la mano percorre una traiettoria che la porta non sul bersaglio, ma alla sua destra. L'*aftereffect* dovuto all'adattamento, che tipicamente scompare abbastanza rapidamente, fornisce informazioni preziose sulla natura delle rappresentazioni coinvolte nella guida del movimento. In particolare, nel suo esperimento Harris chiedeva ai partecipanti di indossare un prisma ottico che alterava l'informazione visiva su una delle due mani (la mano *adattata*), spostandone la proiezione sulla retina. Dopo un numero sufficiente di prove i partecipanti erano in grado di raggiungere correttamente il bersaglio nonostante la distorsione, indicando che l'adattamento era completo. A quel punto, Harris toglieva i prismi e misurava l'*aftereffect*, in tre condizioni diverse e sia per la mano adattata sia per la mano *non adattata*. La prima delle tre condizioni prevedeva che la posizione del bersaglio fosse specificata dalla visione. La seconda, che fosse specificata dal suono di un cicalino fissato nella posizione appropriata sotto il tavolo. La terza condizione, infine, non prevedeva la presentazione di un bersaglio e il partecipante doveva semplicemente muovere il braccio in avanti (vedi tab. 3.1). Nella prima condizione un telo consentiva di vedere il bersaglio ma non la mano in movimento (che era invece ovviamente visibile durante le prove di adattamento). Nelle altre due condizioni il movimento veniva effettuato a occhi chiusi.

Utilizzando questa tecnica, Harris ha osservato che l'*aftereffect* rimaneva sostanzialmente identico in tutte e tre le condizioni, ma solo per la mano adattata. Quando la prova veniva svolta con la mano non adattata, l'*aftereffect* scompariva. Ciò indica che la ricalibrazione della trasformazione sensomotoria, a seguito dell'adattamento alla distorsione visiva, si trasferisce anche ad altre mo-

dalità di senso (infatti si osserva anche l'**«aftereffect» con un bersaglio acustico**) e a una sequenza di attivazioni muscolari non necessariamente identica a quella adattata (infatti si osserva l'*aftereffect* anche con un movimento genericamente in avanti, senza un bersaglio che ne fissi precisamente la traiettoria). Dunque la rappresentazione spaziale che viene ricalibrata non è limitata alla relazione fra visione e movimento e non dipende da una mera forma di apprendimento motorio, in cui il sistema apprende ad associare una nuova configurazione di attivazioni muscolari a un certo stimolo visivo. Per guidare la mano, il sistema sembra invece utilizzare una rappresentazione multisensoriale della posizione del corpo e dello spazio attorno al corpo.

Questa rappresentazione è più complessa e astratta di quella che risulta dal mero confronto fra posizioni della mano e del bersaglio all'interno di un singolo canale sensoriale. Allo stesso tempo, non riguarda la rappresentazione del corpo nel suo complesso ma solo quella dello specifico effettore coinvolto nel movimento (infatti, l'*aftereffect* non si trasferisce alla mano non adattata). Da questo punto di vista, dunque, l'esperimento di Harris porta a conclusioni coerenti con altre riportate in letteratura da Haggard e colleghi [2000] e da van Beers, Wolpert e Haggard [2002] sulla percezione multisensoriale della posizione statica della mano. Inoltre, i risultati di Harris anticipano, e sono in buon accordo, con numerose osservazioni sulla fisiologia dei neuroni in alcuni circuiti neurali che coinvolgono aree dei lobi frontale e parietale del cervello primate. Una larga parte di questi neuroni, infatti, risponde sia a stimolazione visiva sia a stimolazione tattile (**neuroni bimodali**), codificando gli stimoli in coordinate ancorate al corpo (anziché alla retina, come nel caso di un tipico neurone visivo) e rispetto a una molteplicità di sistemi di riferimento somatici, centrati ad esempio sulla mano, sulla testa, sul braccio [Fogassi *et al.* 1996; Gentilucci *et al.* 1988; vedi anche Graziano 2001]. Questo tipo di proprietà funzionali rappresenta, presumibilmente, il sostrato neurale di una rappresentazione astratta del corpo e dello spazio che è indipendente sia dal tipo di canale sensoriale in ingresso sia dal tipo di comando motorio coinvolto [Andersen e Buneo 2002]. Parleremo ancora di questo nel capitolo 10 nella terza parte di questo libro.

3. RAGGIUNGERE E AFFERRARE

Molto spesso lo spostamento della mano non ha solo lo scopo di riposizionare l'arto o di indicare una posizione, ma serve anche a raggiungere un oggetto per afferrarlo. Un aspetto interessante è che le due componenti di questa azione,

lo **spostamento della mano** verso l'oggetto e la **conformazione delle dita** per afferrarlo, non vengono eseguite in sequenza ma vengono invece coordinate in maniera sofisticata. Infatti, la mano non aspetta di arrivare in contatto con l'oggetto per aprirsi, ma inizia ad assumere una forma appropriata *durante* lo spostamento, raggiungendo la sua massima apertura molto prima di raggiungere il bersaglio. In questo senso, dunque, lo spostamento della mano finalizzato al solo raggiungimento non è equivalente, dal punto di vista dell'utilizzo dell'informazione percettiva e della guida del movimento, allo spostamento coordinato con la prensione. Se, infatti, è evidente che le due componenti possono essere eseguite anche indipendentemente l'una dall'altra, quando queste vengono eseguite insieme le loro caratteristiche temporali e spaziali vengono adattate allo scopo di portare la mano nelle vicinanze dell'oggetto preparando contemporaneamente la mano per la manipolazione.

Grazie all'utilizzo di apparecchiature in grado di registrare la posizione di appositi *marker* posizionati sulle dita e sul polso, moltissime ricerche hanno studiato il coordinamento delle due componenti nel tempo e la loro modulazione in funzione delle caratteristiche dell'oggetto da afferrare. Inizialmente, queste ricerche si sono concentrate principalmente sulla cosiddetta **presa di precisione** (*precision grip*), utilizzata per afferrare oggetti relativamente piccoli grazie all'opposizione del pollice e dell'indice (fig. 3.3; per diversi punti di vista sulla cinematica della presa di precisione, vedi Jeannerod [1981]; Smeets e Brenner [1999]; in italiano è disponibile la rassegna di Bonfiglioli [2000]). Ricerche più recenti hanno iniziato a studiare anche la **conformazione globale della mano** utilizzando apparecchiature più sofisticate in cui i partecipanti indossano veri e propri guanti su cui sono inseriti sensori in grado di registrare simultaneamente la traiettoria di tutte le articolazioni.

Complessivamente, queste ricerche hanno evidenziato una serie di principi che regolano l'azione di raggiungimento e prensione. In primo luogo, l'ampiezza

fig. 3.3. Raggiungimento e prensione nella presa di precisione. La mano si solleva durante lo spostamento e le dita si aprono, raggiungendo il punto di massima apertura poco dopo la prima metà del movimento. Sia l'ampiezza della massima apertura sia la durata delle fasi precedente e successiva alla massima apertura dipendono dalle caratteristiche dell'oggetto.

dell'**apertura massima delle dita** (*Maximum Grip Aperture*, o MGA) è sempre maggiore della grandezza fisica dell'oggetto, ma in modo proporzionale a questa (*grip scaling*). Detto altrimenti, per oggetti più grandi l'MGA tende a essere maggiore che per oggetti più piccoli [Marteniuk *et al.* 1987]. Per questo motivo la maggior parte dei ricercatori ritiene che il *grip scaling* fornisca informazioni sulla rappresentazione della grandezza utilizzata nella trasformazione sensomotoria per la guida della prensione. In secondo luogo, il momento in cui avviene l'MGA coincide con l'inizio di una **fase di decelerazione**, la quale tende a essere più lunga quanto più piccolo è l'oggetto da afferrare. L'allungarsi della fase di decelerazione viene generalmente interpretato come una conseguenza della maggiore precisione necessaria per guidare la mano verso un oggetto piccolo utilizzando il segnale visivo afferente disponibile durante il movimento [Marteniuk *et al.* 1990]. Quest'ultima fase del movimento sarebbe dunque l'equivalente, nel raggiungimento e prensione, della fase che Woodworth chiamava monitoraggio, mentre l'accelerazione iniziale della mano e l'apertura delle dita dipenderebbero principalmente da elaborazione a carico della fase di impulso iniziale – un altro dato a favore dell'idea che l'MGA riflette la rappresentazione della grandezza nel sistema che effettua la trasformazione sensomotoria iniziale più che in quello che controlla il movimento durante il suo svolgimento. Infine, la configurazione complessiva della mano viene influenzata, specialmente nella parte finale dell'azione [Santello e Soechting 1998], anche dalla forma dell'oggetto da afferrare. Ad esempio, il direzionamento delle dita verso i punti finali di contatto viene modulato dalla posizione percepita del centro di massa dell'oggetto [Lukos, Ansuini e Santello 2007] e la modifica improvvisa della sua forma durante l'esecuzione determina una rapida riorganizzazione nel direzionamento di tutte le dita [Ansuini *et al.* 2007].

Questi risultati pongono il problema di individuare le fonti di informazione percettiva utilizzate per programmare la prensione nella maniera appropriata. In altre parole, come fa la mano a «sapere» in quale direzione deve muoversi, di quanto deve muoversi, e soprattutto quanto e come aprirsi, senza avere ancora toccato l'oggetto? Le ricerche disponibili suggeriscono che le caratteristiche del movimento di raggiungimento e prensione dipendono in primo luogo da informazioni raccolte attraverso il canale visivo e che specificano la distanza dal corpo, la grandezza e la forma tridimensionale dell'oggetto [per approfondimenti sulla natura di queste informazioni, vedi Bruno 2004b; di Pellegrino 2004]. In particolare, diversi lavori hanno mostrato come l'apertura delle dita nella presa di precisione dipenda in modo cruciale dalla stereopsi, ossia dal sistema in grado di elaborare le piccole differenze nelle immagini proiettate sulle

due retine (disparità retiniche) per ricostruire la struttura tridimensionale di un oggetto e la sua distanza [Bradshaw *et al.* 2004]. Informazioni come la grandezza retinica, l'altezza sul piano retinico, rilevabili anche a livello monoculare, influenzerebbero la presa solo indirettamente, dopo essere state elaborate da meccanismi coinvolti nel riconoscimento. Dati cruciali a favore di questa conclusione provengono da uno studio di Marotta, Behrmann e Goodale [1997] su una paziente con agnosia visiva, una condizione che produce un deficit profondo nella capacità di descrivere la struttura tridimensionale di un oggetto e di denominarlo, lasciando però largamente intatta la capacità di guidare azioni rivolte all'oggetto stesso [Milner *et al.* 1991]. Utilizzando manipolazioni che rimuovevano le disparità retiniche, questo studio ha dimostrato che il *grip scaling* della paziente veniva pesantemente alterato nelle prove in cui le disparità non erano disponibili, mentre rimaneva paragonabile a quello di un individuo senza danno neurologico se queste erano presenti.

Questi risultati sono stati confermati anche da numerosi studi in cui la disponibilità di informazioni binoculari e monoculari sono state manipolate in soggetti sani. Ad esempio, la rimozione delle disparità retiniche modifica il *grip scaling* da parte dei soggetti senza danno neurologico quando l'azione di raggiungimento e prensione viene eseguita a circuito aperto (vedi par. 2), ossia spegnendo la luce quando il movimento ha inizio, mentre in queste condizioni lo spostamento della mano rimane comparabile a quello eseguito a circuito chiuso in cui l'informazione visiva rimane disponibile durante l'esecuzione [Melmoth e Grant 2006; Watt e Bradshaw 2000]. Inoltre, le manipolazioni del campo visivo influenzano in maniera diversa il *grip scaling* e lo spostamento della mano. Quando lo stimolo contiene solo informazioni dalla periferia della retina si osservano alterazioni in entrambe le componenti, mentre quando contiene solo informazioni dal centro della retina si osservano solo alterazioni nella componente di raggiungimento [Sivak e MacKenzie 1990]. Il restringimento progressivo del campo visivo causa una riduzione corrispondente nello spostamento (la mano si ferma prima di arrivare all'oggetto) mentre lascia sostanzialmente intatto il *grip scaling* [Watt, Bradshaw e Rushton 2000]. Questi risultati sono dunque in accordo con l'idea che il *grip scaling* utilizzerebbe prevalentemente informazione binoculare focale e sarebbe relativamente poco influenzato dall'informazione contestuale, cruciale per rilevare le informazioni monoculari come la grandezza e l'altezza retiniche; mentre la componente di raggiungimento utilizzerebbe anche queste ultime.

In conclusione, questi risultati indicano che l'informazione spaziale visiva ha un ruolo importante nella guida del raggiungimento e della prensione, anche se le due componenti dell'azione sembrano differire per quanto riguarda

la maniera in cui questa informazione viene utilizzata. Tutto ciò non esclude tuttavia che il sistema che guida l'azione non sia in grado di sfruttare informazioni rese disponibili da altri meccanismi percettivi. Ad esempio, anche il colore dell'oggetto da afferrare – una proprietà visiva elaborata da meccanismi diversi da quelli della visione spaziale – sembra modulare il *grip scaling* [Gentilucci *et al.* 2001]. Già nella prima metà del secolo scorso, alcuni studi [Warden e Flynn 1926] avevano riportato che, a parità di grandezza fisica, i giudizi di grandezza percepita tendono a essere influenzati dal colore dell'oggetto (ad esempio gli oggetti rossi vengono giudicati lievemente più grandi di quelli verdi). La ragione di questo tipo di effetti non è chiara, ma sembra possibile ipotizzare che il sistema percettivo incorpori una sorta di assunzione *a priori*, legata alla covariazione di colore e grandezza negli oggetti naturali tipici delle nicchie ecologiche in cui si è evoluta la nostra specie. Come dire, se la nostra specie per lunghi periodi è stata esposta a foglie ed erbe statisticamente più piccole di frutti e tronchi, il sistema percettivo potrebbe avere evoluto una sorta di associazione di cui si possono ritrovare le tracce ancora oggi.

Se questa ipotesi è corretta, a maggior ragione si potrebbe prevedere che modulazioni nel *grip scaling* avvengano anche da parte degli odori. Alcuni studi recenti hanno confermato questa ipotesi. In una ricerca di Castiello e colleghi [2006], i partecipanti dovevano afferrare un oggetto mentre veniva presentato un odore. I risultati hanno dimostrato che quando l'oggetto da afferrare era piccolo, come ad esempio una fragola, il *grip scaling* evidenziava un MGA maggiore se l'odore compresente corrispondeva a un oggetto più grande, come ad esempio un'arancia, rispetto alle condizioni in cui non c'era odore o l'odore corrispondeva all'oggetto. Quando l'oggetto era grande, come una pesca, ma l'odore corrispondeva a un oggetto più piccolo, come una mandorla, l'effetto si invertiva. In un secondo lavoro, lo stesso gruppo [Tubaldi *et al.* 2008] ha dimostrato un effetto simile anche per la configurazione globale della mano. Quando la visione e l'olfatto corrispondevano a oggetti diversi, i risultati suggerivano, infatti, un'alterazione delle relazioni fra le diverse dita durante la prensione. Quando invece corrispondevano allo stesso oggetto, le dita mostravano il tipico elevato livello di sinergia e di coordinazione.

Altre ricerche, infine, hanno evidenziato anche interazioni fra la visione e altri canali sensoriali. In un recente lavoro di Patchay, Haggard e Castiello [2006], i partecipanti dovevano afferrare un oggetto posto su un piano mentre con l'altra mano stringevano un altro oggetto posizionato immediatamente sotto il piano. I risultati hanno dimostrato che la cinematica della mano che eseguiva la prensione era influenzata, oltre che dalla grandezza percepita visivamente, anche dalle dimensioni percepite per via tattile dell'oggetto stretto nell'altra mano. Un aspetto interessante di questi risultati è che l'interazione visuotattile era evidente sia quando l'oggetto tattile era posizionato esattamente sotto l'oggetto da afferrare (coincidenza spaziale di tatto e vista), sia quando era spostato lateralmente rispetto a questo. Questi risultati sembrano quindi indicare che la modulazione del *grip scaling* utilizza una rappresentazione integrata delle proprietà dell'oggetto, condivisa da effettori diversi, e indipendente dalla posizione nello spazio. Si tratta dunque di una tipologia di interazione multisensoriale in contrasto sia con quanto avviene in altri fenomeni di integrazione multisensoriale in cui la coincidenza spaziale è critica (come abbiamo già accennato nell'introduzione al libro, e come discuteremo in maggiore dettaglio nel cap. 10), sia con le caratteristiche delle integrazioni multisensoriali deputate alla localizzazione della mano, di cui abbiamo accennato nei paragrafi precedenti per lo spostamento, e di nuovo, per la localizzazione statica nel precedente paragrafo sulle mappe spaziali separate per la mano sinistra e destra.

Da un altro punto di vista tuttavia l'idea che i processi multisensoriali possano, in alcuni casi, utilizzare **schemi di riferimento centrati sull'oggetto** invece che basati sulla coincidenza spaziale potrebbe riflettere un altro tipo di funzione adattiva. Per certe tipologie di azioni complesse la possibilità di condividere rappresentazioni oggettuali fra diverse parti del corpo potrebbe costituire un importante vantaggio. Si pensi ad esempio alle situazioni in cui la manipolazione costituisce la base per l'esplorazione dell'oggetto con due mani, dove ovviamente la capacità di «mettere insieme» l'informazione tattile raccolta dalle due parti del corpo e di integrarle con l'informazione visiva rappresenta un presupposto per la ricostruzione corretta della struttura tridimensionale. O, ancora, si consideri la situazione in cui la prensione dell'oggetto costituisce la fase iniziale di un'azione che culmina con il movimento che lo porta alla bocca per un'ulteriore esplorazione (come fanno i bambini piccoli) o per l'ingestione. Quest'ultima situazione richiede una coordinazione nell'utilizzo dell'informazione visiva, per guidare il movimento di raggiungimento e prensione, e di quella propriocettiva per l'ingestione. Infatti, la registrazione dei movimenti oculari nelle azioni di questo tipo dimostra che gli individui spostano gli occhi sul bersaglio in preparazione allo spostamento della mano finalizzato alla prensione, ma una volta afferratolo disattivano la componente oculomotoria utilizzando le sole informazioni muscolari e propriocettive per portarlo alla bocca [de Bruin *et al.* 2008]. In questo processo, una **condivisione della rappresentazione dell'oggetto** fra il sistema che guida la mano e quello che controlla la bocca avrebbe degli evidenti vantaggi.

In accordo con questa aspettativa, diverse ricerche hanno dimostrato che la prensione di un oggetto e la conformazione propriocettiva della bocca si influenzano a vicenda. Ad esempio, Castiello [1997] ha studiato la prestazione di un gruppo di partecipanti che dovevano afferrare e portare alla bocca dei pezzi di formaggio di diverse dimensioni. I risultati hanno evidenziato che, come avviene per il *grip scaling*, l'apertura della bocca era proporzionale alle dimensioni dell'oggetto. Inoltre, sempre come avviene per il *grip scaling*, la massima apertura della bocca tendeva ad avvenire prima per i pezzi piccoli rispetto ai pezzi grandi. Questo risultato è in accordo con le proprietà funzionali di una classe di neuroni nella corteccia premotoria (area F5) della scimmia, che ha evidenziato [Matelli, Luppino e Rizzolatti 1985] come questi neuroni tendano ad attivarsi in maniera simile sia quando l'animale afferra un oggetto con la mano sia quando l'afferramento avviene con la bocca. Inoltre esperimenti che hanno studiato il *grip scaling* quando i partecipanti dovevano contemporaneamente aprire la bocca hanno dimostrato che questa seconda azione, anche se non specificamente finalizzata all'ingestione, veniva influenzata dalla grandezza dell'oggetto in maniera simile all'effetto sul *grip scaling* [Gentilucci *et al.* 2001]. Queste ricerche confermano quindi che azioni coordinate che coinvolgono effettori diversi possono utilizzare rappresentazioni comuni delle caratteristiche dell'oggetto.

In una prospettiva multisensoriale, questi risultati convergono nel suggerire che l'azione di raggiungimento e prensione coinvolga una molteplicità di rappresentazioni multisensoriali dello spazio e degli oggetti, di cui stiamo appena cominciando a capire le caratteristiche. Sembra plausibile, ad esempio, che anche le informazioni legate al suono possano venire utilizzate nella guida di questo tipo di azioni, e che la combinazione dei canali visivo, propriocettivo, olfattivo e acustico possa seguire regole diverse a seconda della finalità dell'azione stessa (ad esempio, esplorativa o finalizzata all'ingestione) e delle informazioni disponibili nei diversi canali [Zahariev e MacKenzie 2007]. Sarà compito della ricerca futura chiarire la natura di queste regole e delle strutture neurali coinvolte nella loro implementazione.

4. CAMMINARE E MANTENERSI IN EQUILIBRIO

La percezione non è coinvolta solo nella guida delle azioni manuali, ma entra in gioco anche in tutte le altre situazioni in cui muoviamo il corpo. Fra queste, è particolarmente interessante il coinvolgimento dei processi percettivi

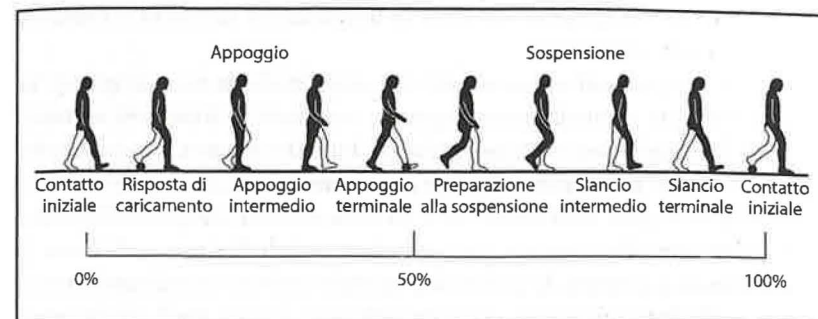



fig. 3.4. Il ciclo del passo nella deambulazione. Ogni passo si suddivide in una fase di appoggio, in cui il piede tocca terra, e in una fase di sospensione o oscillazione in cui il piede viene lanciato in avanti mentre il corpo si sposta anch'esso in avanti.

multisensoriali nella guida della **deambulazione** e nel concomitante controllo della **postura**. L'esame di questi due altri processi motori ci consente di occuparci di un altro aspetto della complessa interazione fra percezione corporea e movimento, vale a dire, il ruolo dei processi multisensoriali nella modulazione di comportamenti generati a livello sottocorticale. L'organizzazione ritmica della deambulazione è illustrata nella figura 3.4, che riassume le due fasi del ciclo di un passo. Come si vede nella figura, ogni passo può essere pensato come una sorta di caduta controllata, in cui il peso corporeo viene trasferito dal piede di appoggio all'altro piede, a seguito dello slancio di questo in avanti, mentre il baricentro del corpo si sposta anch'esso in avanti.

Secondo la concezione tradizionale, la configurazione spaziotemporale dei movimenti degli arti durante la locomozione è controllata fondamentalmente da attività riflesse [Easton 1972]. Affermare che la deambulazione e la postura sono basate su **riflessi** equivale a dire che sono comportamenti in cui l'organismo reagisce alla stimolazione periferica mediante configurazioni motorie automatiche, determinate geneticamente e gestite da circuiti neurali che non coinvolgono processi corticali superiori. A sostegno del ruolo di processi automatici determinati geneticamente nel produrre risposte motorie, va detto che riflessi di questo tipo sono facilmente osservabili nei neonati. Inoltre diversi studi hanno dimostrato una chiara periodicità nell'attività registrabile a livello spinale in modelli animali [Grillner 1975; Marder e Calabrese 1996]. Ad esempio, in esperimenti su gatti a cui venivano interrotte sia le connessioni efferenti fra corteccia e corda spinale, sia le connessioni afferenti dai muscoli, è stato possibile osservare che a seguito di stimolazione elettrica le fibre efferenti della corda spinale esibiscono

un'attività ritmica nonostante l'assenza di informazioni sensoriali o comandi motori dalla corteccia.

Queste osservazioni sono coerenti con una circuiteria neurale spinale in grado di produrre autonomamente sequenze complesse di attivazioni motorie, comparabili a quelle necessarie per produrre un ciclo di passo. Questi circuiti vengono chiamati **generatori di configurazioni motorie** (*pattern generator*). In altri esperimenti, gatti a cui erano recise sia le vie efferenti sia quelle afferenti a livello del mesencefalo venivano posti su pedane mobili. Nel momento in cui la pedana iniziava a scorrere, le gambe dell'animale venivano inizialmente trascinate ma poco dopo iniziavano a eseguire sequenze di passi simili alla locomozione naturale. Inoltre all'aumentare della velocità i passi diventavano più veloci e potevano diventare movimenti di corsa alle velocità più elevate. Dunque i segnali provenienti dai muscoli delle gambe sono sufficienti per attivare e modulare l'attività riflessa che controlla il passo, anche in assenza di comandi motori intenzionali da parte dell'animale cosciente (infatti, questi segnali afferenti non raggiungono la corteccia a causa dell'operazione di deafferentazione).

Presi nel loro complesso, questi risultati sono coerenti con un **modello a due componenti**: dei *pattern generator*, a livello spinale, in grado di produrre il ciclo del passo anche in assenza di segnali sensoriali; e dei meccanismi riflessi, controllati a livello cerebellare, in grado di attivare e modulare la generazione dei comandi motori in funzione dei segnali afferenti dai muscoli. Viene generalmente ritenuto che queste due componenti consentirebbero un controllo della deambulazione particolarmente efficiente, perché basato su una configurazione di comandi motori automatici, ma in grado di venire modificata all'occorrenza da segnali muscolari [Schmidt e Lee 1999], ad esempio quando questi segnalano mutamenti inaspettati nell'ambiente.

Meccanismi riflessi sono stati studiati anche nel caso del **controllo della postura** nell'organismo statico. In una serie di classici esperimenti, Nashner [1976] poneva i partecipanti in piedi su una piattaforma che poteva scivolare o ruotare all'indietro (vedi fig. 3.5). Entrambe le condizioni comportano lo stiramento nel muscolo gastrocnemio (gemelli) evocando un riflesso che ha l'effetto di aprire l'articolazione della caviglia e di spostare il peso all'indietro. Nella situazione dell'esperimento di Nashner, tale spostamento del peso all'indietro ha l'effetto di stabilizzare la postura quando la piattaforma scivola all'indietro, perché in questo caso il movimento sbilancia il corpo in avanti. Quando invece la piattaforma ruota, il peso del corpo viene portato all'indietro, non in avanti, e quindi il riflesso di stiramento causa una destabilizzazione ancora maggiore della postura. Studiando l'evoluzione del riflesso nel corso di prove ripetute,

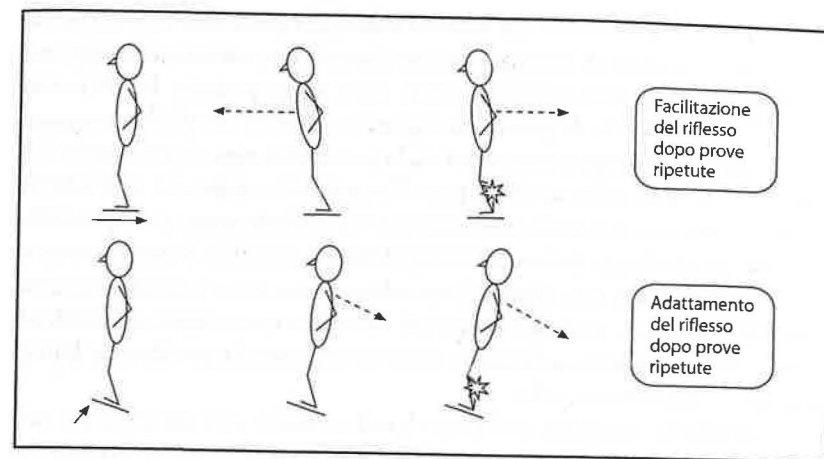



fig. 3.5. Modulazione del riflesso di stiramento del muscolo gastrocnemio in funzione delle conseguenze sulla postura. Sia lo scivolamento sia la rotazione della piattaforma su cui è posto il partecipante (sinistra) evocano il riflesso (destra). Poiché le due situazioni hanno conseguenze diverse sulla postura (centro), il riflesso la stabilizza quando il corpo viene sbilanciato in avanti ma non quando viene sbilanciato all'indietro (destra). Dopo una serie di prove si osserva facilitazione del riflesso nel primo caso e adattamento del riflesso nel secondo.

fonte: Adattata da Nashner [1976].

Nashner ha osservato che l'attività elettromiografica al livello del muscolo tendeva ad aumentare (facilitazione del riflesso) nel caso in cui l'esito era la stabilizzazione della postura, mentre tendeva a diminuire (adattamento del riflesso) quando l'esito era opposto. Inoltre, la modulazione del riflesso in funzione degli esiti sulla postura era assente in pazienti con patologie del cervelletto. I risultati di Nashner confermano dunque l'esistenza di circuiti riflessi per la postura e la deambulazione, controllati da meccanismi cerebellari in grado di modularli.

Le due situazioni discusse nei paragrafi precedenti suggeriscono dunque che la deambulazione e la postura sfruttano informazioni posturali e cinetiche provenienti dai recettori nei muscoli nel contesto di circuiti riflessi controllati a livello del cervelletto. Ciò non significa però che tutta la deambulazione e il controllo dell'equilibrio siano basati solo su questi meccanismi. Molte altre ricerche hanno mostrato, infatti, che in compiti più complessi o più ricchi dal punto di vista informativo la locomozione e la postura vengono influenzate profondamente da **processi multisensoriali superiori** che tengono conto di segnali efferenti elaborati da precisi centri corticali. Un confronto interessante con gli esperimenti di Nashner appena descritti è il cosiddetto esperimento della stanza

mobile [Lee e Lishman 1975]. La stanza mobile è una specie di grande corridoio con un soffitto a forma di rettangolo stretto e lungo, due pareti laterali anch'esse rettangolari, e una parete frontale di forma più o meno quadrata. L'altra parete corta non c'è, in modo da poter fare entrare i volontari che si sottopongono all'esperimento. Ma soprattutto, e qui sta la peculiarità, non c'è pavimento. La stanza mobile viene sollevata di un paio di centimetri e appesa al vero soffitto del laboratorio, con dei giunti che consentono di traslarla avanti e indietro nella direzione del lato lungo. Infine, all'interno della stanza mobile le pseudopareti e lo pseudosoffitto non sono dipinti di un colore omogeneo, ma possiedono una ricca struttura ottica – ad esempio, macchie disposte a caso o figure geometriche irregolari di vari colori, o anche semplicemente una carta da parati un po' kitsch di quelle che usavano una volta.

L'esperimento funziona così: si entra nella stanza e ci si sistema con i talloni su un piccolo gradino, circa a metà della lunghezza, in modo da essere in equilibrio un po' meno stabile del solito (questo non è indispensabile ma rende l'effetto più drammatico). Il compito consiste semplicemente nel guardare verso la parete stretta in fondo. Quello che fa lo sperimentatore è muovere la stanza nella direzione della linea di mira. Il risultato di questa semplice manipolazione è sorprendente: i volontari perdono l'equilibrio e tendono a cadere all'indietro; alcuni talvolta finiscono davvero col sedere per terra (fig. 3.6, destra). Se il controllo della postura fosse basato solo su riflessi evocati dalle informazioni provenienti dai muscoli, una cosa del genere non dovrebbe verificarsi. Infatti, la stanza non è posata sul pavimento, mentre i piedi dei volontari lo sono. Quindi quando la stanza viene spostata, a differenza di quanto accade nel paradigma di Nashner, non arriva nessuno stimolo ai recettori nei muscoli delle gambe (o di qualsiasi altro distretto corporeo). L'informazione che produce la perturbazione della postura è invece di natura *visiva*.

Per capire di cosa si tratta, considerate quello che avviene su un piano di proiezione ottica, come la retina, quando un organismo si sposta rispetto a oggetti ambientali statici. In primo luogo, dato che la posizione della retina cambia rispetto all'ambiente tridimensionale, anche la proiezione si sposta in accordo con i principi della geometria proiettiva. Ad esempio se la fissazione è nella direzione del movimento, c'è un punto speciale sulla retina in cui non si verifica alcuno spostamento (quello nella direzione della fissazione), e tutti gli elementi circostanti a quel punto si muovono, sulla retina, allontanandosi dal punto fisso in una direzione radiale (fig. 3.6, sinistra). Questo tipo di movimento radiale sulla retina è un esempio di **flusso ottico** e il punto singolare che non si muove in alcuna direzione retinica viene chiamato **fuoco di espansione**.

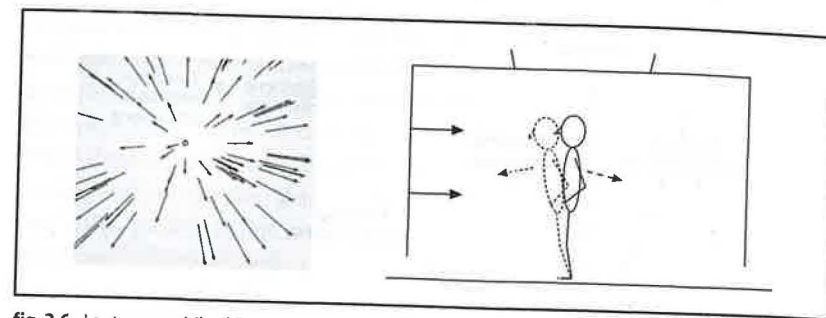
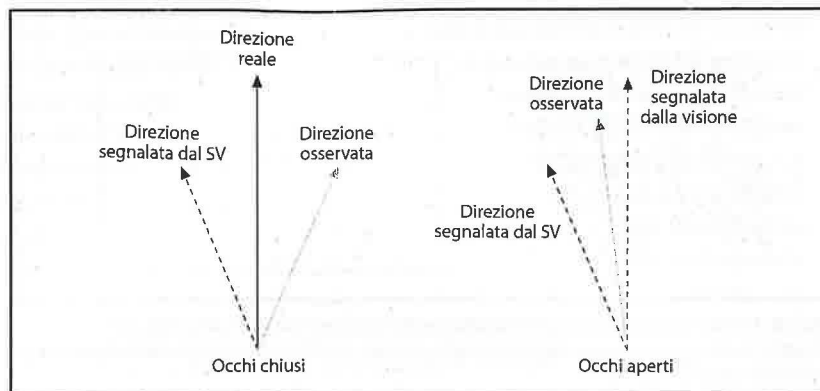



fig. 3.6. La stanza mobile. Il flusso ottico radiale specifica movimento in avanti del corpo anche in assenza di segnali muscolari e vestibolari. Il processo multisensoriale che influenza la postura risponde spostando il peso del corpo all'indietro.

Una grande quantità di ricerche ha dimostrato che noi, ma anche molti animali non umani, siamo sensibili all'informazione spaziale contenuta nel flusso ottico e la utilizziamo per controllare la postura e la direzione della locomozione. Esperimenti psicofisici suggeriscono che le informazioni contenute nel flusso ottico sono una condizione sufficiente, anche se non necessaria, per percepire la **direzione della propria locomozione** (il cosiddetto *heading* – la direzione in cui si punta, la rotta). Specificamente, i giudizi di *heading* sono modulati dalla posizione del fuoco di espansione nel flusso ottico [Bruggeman, Zosh e Warren 2007; Prokop, Schubert e Berger 1997; Warren *et al.* 2001] quando questo è presente, ma le persone sono in grado di usare anche altre informazioni quando il flusso ottico non è disponibile [Loomis *et al.* 2006; Macuga *et al.* 2006]. L'inverso del tasso di espansione radiale (il cosiddetto parametro *tau*) controlla la percezione del tempo che manca per il contatto con l'oggetto – o la parete – verso cui ci stiamo muovendo [Yilmaz e Warren 1995; Lee 1976]. Inoltre esperimenti su scimmie [Britten e van Wezel 1998] hanno indicato che le strutture neurali coinvolte nell'analisi del flusso ottico potrebbero essere collocate nell'area temporale mediale superiore (MST, *Medial Superior Temporal area*), un'area specializzata per la percezione del movimento. Studi di risonanza magnetica funzionale dimostrano che l'area considerata omologa ad MST nel cervello dell'uomo risponde in maniera selettiva al flusso ottico radiale [Morrone *et al.* 2000]. Nel caso della stanza mobile, il movimento della stanza ha proprio l'effetto di produrre un flusso ottico radiale compatibile con il movimento in avanti dell'osservatore, a cui il partecipante reagisce portando il corpo all'indietro. I risultati dimostrano quindi che le risposte muscolari di stabilizzazione della postura (anche se, ovviamente, nella stanza mobile l'effetto è in realtà de-

fig. 3.7. Errori sistematici nella direzione della propria locomozione (*heading*) da parte di un paziente che ha subito NVU a destra. La differenza fra le afferenze vestibolari indica rotazione verso sinistra. Pertanto a occhi chiusi il paziente compensa virando verso destra. A occhi aperti lo *heading* è invece un compromesso fra l'informazione vestibolare e quella visiva.

stabilizzante) possono essere controllate anche da informazioni visive elaborate a livello corticale e non solo dai meccanismi riflessi descritti prima.

Gli effetti osservabili con il **paradigma della stanza mobile** indicano che anche nella deambulazione e nel controllo della postura la visione ha un ruolo importante. Infatti, anche in presenza di informazioni contraddittorie dai muscoli e dal sistema vestibolare, il flusso ottico può essere sufficiente per indurre l'impressione di movimento del proprio corpo. Succede nella stanza mobile, ma anche, più comunemente, negli scompartimenti dei treni fermi al binario. A chi non è successo di pensare: «Finalmente il treno sta partendo!» e di accorgersi, solo un secondo dopo, che non era il nostro treno a muoversi, ma il treno nel binario accanto. Visto attraverso il finestrino, il flusso ottico (in questo caso non radiale, ma «lamellare») prodotto dal treno *in arrivo* alla stazione mentre siamo fermi è equivalente a quello che si produrrebbe sulla nostra retina quando il nostro treno parte e accanto ci sono dei vagoni fermi. Se sussistono certe condizioni, questo tipo di stimolo visivo può essere sufficiente a indurre un movimento apparente del Sé molto convincente, anche se di breve durata. In condizioni normali, le informazioni visive interagiscono con quelle propriocettive-muscolari e vestibolari per guidare la locomozione e controllare la postura.

Un'osservazione istruttiva a questo proposito riguarda lo studio della deambulazione nei pazienti trattati con **neurotomia vestibolare unilaterale** (NVU) a seguito di diagnosi di malattia di Menière, una grave labirintopatia che causa

vertigini invalidanti e disturbi dell'udito. La resezione delle fibre afferenti dal vestibolo affetto può aiutare a eliminare i disturbi uditivi e le vertigini, al costo di un'alterazione nell'informazione vestibolare a cui i pazienti riescono ad adattarsi solo in parte. Uno studio su pazienti trattati con NVU [Borel *et al.* 2004] dimostra che questi pazienti non riescono a camminare in linea retta a occhi chiusi, ma deviano sistematicamente nella direzione dell'orecchio trattato con NVU. Inoltre, se devono camminare a occhi aperti tendono a deviare nella direzione *opposta* a quella dell'orecchio trattato. Entrambi i risultati sono in accordo con un processo multisensoriale per il controllo della direzione della propria locomozione (fig. 4.9). Il sistema vestibolare, infatti, integra informazioni bilaterali dai vestiboli per codificare gli spostamenti del corpo nello spazio. In particolare, la rotazione del corpo verso una direzione (ad esempio, a sinistra) aumenta la frequenza di scarica nel vestibolo di sinistra e la riduce in quello di destra. Ciò significa che una differenza nei segnali afferenti dai due vestiboli segnala che il corpo ruota verso il lato da cui proviene il segnale più forte. A sua volta, questo implica che quando un paziente trattato con NVU cammina a occhi chiusi, la differenza nei segnali afferenti segnala sempre rotazione verso il lato opposto a quello trattato, dove il segnale è zero e quindi per definizione sempre minore dell'altro. Se la consegna è camminare in linea retta ecco allora che il paziente cercherà di compensare deviando nella direzione opposta, ossia *verso* il lato trattato. Quando cammina a occhi aperti, invece, la sua direzione sarà un compromesso fra la direzione indicata dalla visione e quella indicata dalla differenza fra le afferenze vestibolari. Il risultato sarà dunque una deviazione, meno marcata, nella direzione opposta a quella dell'orecchio trattato.

Questi risultati indicano che il controllo della direzione della propria locomozione dipende dall'elaborazione centrale di informazioni sia visive sia vestibolari. Altri dati interessanti sulla natura multisensoriale della locomozione provengono da **esperimenti con pedane mobili**. In uno di questi lavori [Varraine, Bonnard e Pailhous 2002], i partecipanti camminavano su una pedana posta davanti a un ampio schermo su cui venivano proiettate animazioni coerenti con il flusso ottico di un movimento in avanti del partecipante. Sia le caratteristiche di questo flusso ottico sia la resistenza della pedana erano controllate da computer e potevano essere modificate durante il movimento. Misurando le variazioni di velocità e di spinta propulsiva successive a tali modifiche, i ricercatori hanno dimostrato che queste variazioni erano più efficaci e avevano tempistiche diverse se le modifiche avvenivano in entrambi i canali rispetto alle risposte a modifiche sulle sole caratteristiche ottiche o meccaniche. In un altro lavoro, Anstis [1995] ha esaminato un curioso *aftereffect* successivo alla corsa prolungata su pedana

mobile. Come sanno molti frequentatori di palestre, la corsa sul *tapis roulant* da allenamento al chiuso richiede inizialmente un po' di attenzione. Che le prime esperienze con queste attrezzature pongano dei problemi al coordinamento di movimento e postura dipende dal fatto che l'apparecchio ci consente di effettuare il movimento di corsa in modo pressoché normale, ma restando fermi in un punto. Durante la corsa sulla pedana dunque noi riceviamo informazioni muscolari coerenti con il movimento in avanti, mentre la visione invia informazioni coerenti con l'assenza di movimento. Tuttavia, dopo pochi minuti tutti imparano a correre sulla pedana senza grandi difficoltà, anche se l'adattamento alla stimolazione multisensoriale inusuale non è privo di conseguenze. Una prova molto semplice può essere effettuata se si dispone di una pedana con l'inclinazione regolabile. Basta regolare la pedana su una salita abbastanza faticosa e camminare o correre per una decina di minuti; quindi riportare velocemente la pedana nella posizione orizzontale. Si avrà l'impressione di camminare in leggera discesa.

L'*aftereffect* studiato da Anstis si riferisce a ciò che accade quando si cerca di correre sul posto a occhi chiusi. Prima di correre sulla pedana mobile, i partecipanti non hanno alcuna difficoltà a eseguire questo compito correttamente, ossia restando effettivamente fermi sul posto. Dopo la corsa prolungata sulla pedana invece qualcosa cambia, ed eseguire il compito diventa difficile. Quasi tutti hanno l'impressione di correre sul posto quando in realtà stanno spostandosi leggermente in avanti. Questo accade sia se la corsa di adattamento viene fatta a occhi chiusi sia se viene fatta a occhi aperti [Durgin e Pelah 1998]. Anstis ha dimostrato inoltre che l'adattamento che causa questo *aftereffect* non si trasferisce da un arto all'altro. Se invece di correre si chiede ai partecipanti di saltellare su una gamba sola, l'*aftereffect* si osserva quando questi saltellano sul posto a occhi chiusi con la *stessa* gamba usata durante la corsa sulla pedana, ma non con l'altra [vedi anche Durgin, Fox e Hoon Kim 2003]. L'adattamento visuomotorio che entra in gioco in queste condizioni sembra avere dunque delle caratteristiche in comune con i fenomeni osservati da Harris [1963] relativamente alla guida multisensoriale dello spostamento della mano (vedi par. 3.2 e tab. 3.1). Infatti, sarebbe causato da processi multisensoriali di ricalibrazione della propriocezione muscolare in funzione dell'informazione visiva. Pensiamo a cosa accade durante la corsa sulla pedana mobile. Dato che la posizione del corridore rispetto all'ambiente rimane costante, il flusso ottico è nullo. Dunque l'informazione muscolare che precedentemente era associata a un flusso ottico che specifica corsa in avanti viene ora associata all'assenza di flusso ottico. In questa maniera, la calibratura delle afferenze propriocettive (le sensazioni di

pressione e scivolamento rispetto alla superficie sotto i piedi) viene alterata e il punto corrispondente alla percezione di zero movimento corrisponde in realtà a un certo spostamento in avanti reale [Durgin e Pelah 1999].

In conclusione, i dati sperimentali indicano dunque che locomozione e postura sono influenzate da processi che combinano segnali multisensoriali in modo sofisticato. I segnali coinvolti provengono dalla visione, dal sistema vestibolare, dai segnali propriocettivi provenienti dai muscoli, da quelli tattili relativi alle modalità di contatto con le superfici di supporto. Inoltre le modalità della loro combinazione tengono conto anche dei modelli interni ed efferenti relativi ai movimenti in corso. A questo proposito un apparato piuttosto semplice da realizzare, ma estremamente istruttivo, è il cosiddetto tamburo ottico-cinetico. Il **tamburo ottico-cinetico** è una pedana circolare in grado di ruotare attorno a una barra centrale fissa dove viene montato un corrimano. Attorno alla pedana c'è una superficie cilindrica in grado anch'essa di ruotare. L'apparato consente quindi di variare indipendentemente il movimento delle pareti (controllando il flusso ottico) e della pedana (controllando le informazioni tattili e propriocettive). La cosa interessante è che il partecipante non percepisce necessariamente questi movimenti. Al contrario, se la pedana e la parete cilindrica ruotano nella direzione opposta alla direzione del partecipante, e questi effettua dei passi in avanti sostenendosi sul corrimano, il partecipante percepisce che sta camminando in avanti lungo una traiettoria circolare, che sta spingendo il corrimano davanti a sé, e che sia la pedana sia la parete sono ferme [vedi Lackner e DiZio 2000]. La soluzione alternativa, ossia percepire pedana e parete in movimento e sé fermo in un punto, è praticamente impossibile pur essendo ugualmente giustificata dai dati sensoriali. L'effetto suggerisce quindi che l'attribuzione dei movimenti sulla retina allo spostamento dell'ambiente o del soggetto tiene conto della traiettoria e della forza propulsiva dei passi, rappresentata internamente, per generare una sorta di «aspettativa» di flusso ottico afferente coerente con la marcia in avanti. Vale la pena notare, fra l'altro, che qualcosa di simile accade anche nell'illusione del treno fermo in stazione. Anche in quel caso, infatti, l'emergere alla coscienza della percezione di movimento del Sé comporta che questa interpretazione del flusso ottico venga privilegiata, a scapito della sua alternativa ugualmente coerente con la stimolazione disponibile – l'arrivo in stazione di un treno sul binario accanto. Non conosciamo studi in proposito, ma sembra plausibile che fra le condizioni che favoriscono l'insorgere dell'illusione di movimento del Sé in questo caso ci sia proprio l'aspettativa che il proprio treno inizi a muoversi. Come vedremo nel resto del capitolo, infatti, l'idea secondo la quale la percezione e la guida del movimento utilizzano modelli interni

che sostanzialmente «predicono» l'informazione multisensoriale afferente trova supporto anche da altri risultati più specifici.

5. POSTURA E MOVIMENTO IN PAZIENTI DEAFFERENTATI

Supponete di poter scegliere: preferireste perdere la vista o le sensazioni somatiche? La quasi totalità delle persone, poste davanti a un esperimento mentale analogo a questo [Handel 1993], sceglie di conservare la vista. L'unanimità della scelta è giustificata dall'importanza della visione per la comunicazione, il riconoscimento e il movimento nello spazio. Ma come si vive senza sensibilità somatica? Presumibilmente, la facilità con cui le persone sembrerebbero disposte a rinunciare alla sensibilità somatica dipende anche dalla scarsa consapevolezza dell'importanza delle informazioni che provengono dai recettori nella pelle e nei muscoli per un gran numero di processi percettivi di base, in particolare per il controllo del movimento e della postura, e dal fatto che queste funzioni sono in parte automatiche e non richiedono controllo cosciente. Qui ci proponiamo di correggere questo pregiudizio, occupandoci di patologie neurologiche che possono danneggiare permanentemente le fibre nervose che connettono i recettori periferici nella pelle e nei muscoli con il midollo spinale (**deafferentazione somatosensoriale**). I pazienti che soffrono di queste patologie, piuttosto rari, offrono un'opportunità unica di studiare il ruolo delle afferenze somatosensoriali nella percezione multisensoriale del corpo. I lavori che hanno studiato pazienti di questo tipo consentono di trarre due conclusioni principali, entrambe rilevanti per i nostri scopi in questo libro. La prima riguarda specificamente il ruolo dell'informazione tattile e somestesica e la possibilità di sostituirla con le informazioni visive. La seconda, cui abbiamo accennato alla fine del paragrafo precedente, il ruolo di processi interni di rappresentazione del movimento e di costruzione di modelli predittivi dell'afferenza sensoriale.

Riguardo alla prima di queste due conclusioni, le osservazioni più dettagliate sono state effettuate sul paziente I.W. [Cole 1995]. A causa di una grave neuropatia contratta all'età di 19 anni, I.W. ha perso completamente la funzionalità delle fibre nervose mielinate più grandi al livello delle radici dorsali della colonna vertebrale. Queste fibre nervose connettono i fusi neuromuscolari, i recettori tendinei e i recettori cutanei con la corda spinale ed hanno quindi il ruolo di fornire informazioni sulla posizione degli arti, lo stiramento dei muscoli e gli eventi meccanici (pressioni leggere, vibrazioni, stiramento) al livello della pelle. La malattia ha risparmiato invece le fibre nervose che portano informazioni sulla pressione profonda, caldo e freddo, dolore, e le sensazioni muscolari di affaticamento e tensione. Inoltre è utile ricordare che le fibre efferenti che inviano i comandi motori ai muscoli erano anche intatte. Nel celebre saggio *Pride and a daily marathon* il neurofisiologo inglese Jonathan Cole descrive in dettaglio le conseguenze della malattia sulla capacità iniziale di svolgere vita autonoma da parte di I.W., la faticosa riabilitazione e l'eccezionale recupero di molte funzioni percettive e motorie grazie alla ristrutturazione di alcuni processi multisensoriali. All'insorgere della malattia, I.W. era incapace di vita autonoma: non era in grado di mantenere la stazione eretta, di camminare, o di controllare il movimento delle mani pur essendo in grado di muoverle. Disteso a letto a occhi chiusi non aveva idea della posizione dei propri arti e non aveva alcuna sensazione corporea al di sotto del collo.

Grazie agli sforzi di alcuni fisioterapisti e a una forte motivazione, presumibilmente dovuta all'età ancora giovane, nel corso di circa tre anni I.W. è stato in grado di apprendere nuovamente a stare in piedi, camminare, e in generale gestire autonomamente la propria vita. Va chiarito fin dall'inizio, tuttavia, che questo straordinario recupero non è legato a una guarigione del deficit sensitivo. I.W. continua a non avere alcuna sensazione corporea, ma ha imparato nuovamente a usare gli altri sensi in modo diverso per compensare il deficit. Le strategie da lui utilizzate per questo scopo sono istruttive. Per mantenere la postura e camminare, ad esempio, I.W. deve mantenere monitorato costantemente il proprio corpo sfruttando la visione e le informazioni vestibolari. Il suo passo non è fluido e viene effettuato con le ginocchia rigide. Per le azioni che richiedono il controllo delle mani, come scrivere, afferrare e utilizzare oggetti, I.W. utilizza sempre informazioni visive in congiunzione con le sensazioni termiche associate al contatto con le superfici. Ad esempio, quando stira una camicia si rende conto che la mano è arrivata alla fine dell'asse da stiro non solo perché la vede ma anche perché percepisce la variazione di calore. Dati sperimentali sulla cinematica della prensione da parte di un altro paziente [Gentilucci *et al.* 1994] suggeriscono che l'afferramento di un oggetto in presenza di deafferentazione è costituito da una fase iniziale normale, basata sulla programmazione efferente del movimento effettuata per l'appunto su base visiva, e da una fase finale profondamente alterata, più lunga del normale e con aggiustamenti fini più numerosi. Questi dati e le osservazioni su I.W. indicano dunque che l'informazione tattile ha un ruolo importante in questa seconda fase (il monitoraggio di Woodworth), ma può in parte essere vicariata dalla visione focale e da altre sensazioni corporee come quelle legate alla temperatura. Tutto questo è possibile tuttavia solo al prezzo di una fortissima concentrazione

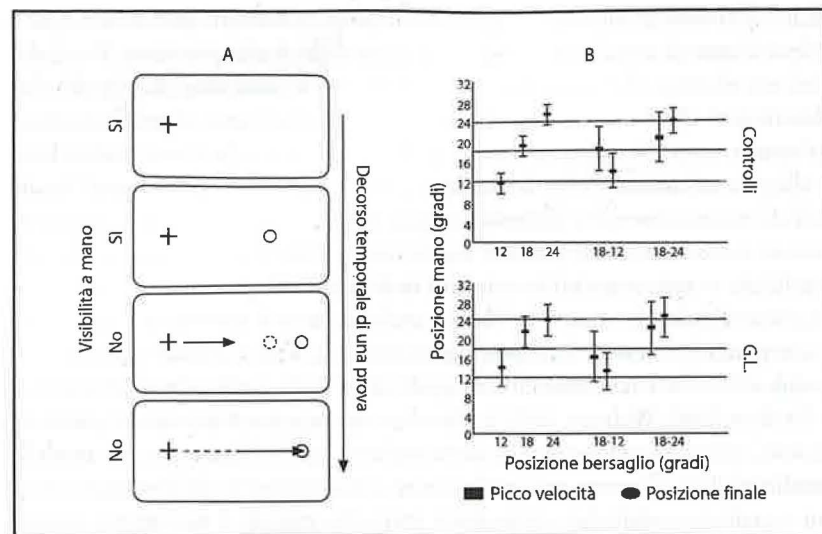
e della disponibilità di tali informazioni alternative. Infatti, I.W. non è in grado di stare in piedi al buio; preferisce dormire con la luce accesa perché risvegliandosi al buio non saprebbe la posizione dei propri arti; quando cammina, qualsiasi distrazione, uno starnuto, l'urto di un passante, lo mette in difficoltà; quando afferra un bicchiere di plastica deve farlo con grandissima cautela perché in assenza di informazioni dalle dita deve monitorare le deformazioni visibili della plastica e rischia di stringere in maniera eccessiva prima di capire che deve fermarsi.

Tuttavia, e sorprendentemente, I.W. è capace di stimare il peso di un oggetto quando lo solleva. Ad esempio, è in grado di dire se una valigia è vuota, mezza piena o piena, anche se non sa dire su che base («Come lo so? Non ne ho idea... lo so e basta») [Cole 1995, 133]. Presumibilmente, I.W. utilizza informazioni visive sulla cinematica dei propri movimenti per farsi un'idea dello sforzo necessario a sollevare l'oggetto. La capacità di utilizzare informazioni visive sul movimento degli arti da parte di un altro organismo per ricavare informazione sullo sforzo impiegato e quindi sulle forze (pesi, attriti) coinvolte nell'azione viene chiamata **«kinematic specification of dynamics»** (KSD). La KSD è una capacità ben documentata in individui non patologici cui vengono presentati filmati contenenti informazioni cinematiche relative ad azioni come il sollevamento di un oggetto [Runeson e Frykholm 1981] (vedi anche la seconda parte di questo libro sulla percezione multisensoriale degli oggetti). Nel caso di I.W., lo stesso principio potrebbe venire applicato anche alla percezione del proprio corpo e non solo ai movimenti compiuti da un'altra persona, integrando tali informazioni visive con informazioni somatestesiche residue, come quelle legate al senso di affaticamento e tensione, e con informazioni vestibolari relative alle variazioni di posizione del corpo rispetto alla gravità. Questa ipotesi è confermata da uno studio sulla discriminazione di pesi da parte di un'altra paziente deafferentata [Fleury *et al.* 1995]. La capacità di discriminare pesi da parte di questa paziente è, infatti, simile a quella dei controlli se può vedersi le braccia, ma decade drasticamente (anche se non completamente) se effettua la pesata senza vedere. Inoltre il picco di velocità nel sollevamento, un indice cinematico, correla molto bene con il peso percepito da parte della paziente, ma non da parte dei controlli. Questi risultati suggeriscono dunque che i pazienti deafferentati in alcuni casi potrebbero avere sviluppato maniere particolari di muovere gli arti, all'apparenza meno fluide ed efficienti, ma utili allo scopo di massimizzare la disponibilità di indici cinematici visivi per vicariare le informazioni muscolari perdute.

Un altro aspetto interessante dei movimenti corporei in questi pazienti è che i processi coinvolti nella **fase preparatoria del movimento**, ossia l'insieme

delle rappresentazioni efferenti usate per la programmazione dell'azione e per i comandi motori iniziali, non sono danneggiati dalla deafferentazione. Per molti anni si è ritenuto che questi processi si riferissero sostanzialmente a quello che Woodworth chiamava impulso iniziale e che oggi chiamiamo controllo a circuito aperto, ossia alla specifica dei comandi motori *prima* che il movimento inizi e alla sua esecuzione in modo ballistico, senza ulteriori aggiustamenti basati sull'afferenza sensoriale. Tuttavia la logica non esclude che queste rappresentazioni motorie efferenti possano anche essere utilizzate per operare ulteriori modifiche o aggiustamenti ai comandi motori iniziali, da fare immediatamente prima di iniziare a muovere l'arto o anche *durante* il movimento stesso. Ma come potrebbero essere fatti questi aggiustamenti, se non usando i segnali sensoriali afferenti? Una possibilità che gode di un certo credito [vedi Desmurget e Grafton 2000; Wolpert 1997] è che oltre alla rappresentazione dei comandi motori, questi processi efferenti siano anche in grado di generare dei **modelli predittivi dell'afferenza sensoriale**, ossia sostanzialmente delle «aspettative» sui segnali sensoriali che arriveranno, una volta iniziato il movimento stesso. Questi modelli predittivi dell'afferenza sensoriale potrebbero essere utilizzati, assieme al modello interno del movimento da eseguire, per costruire una sorta di simulazione interna, in cui dei segnali di errore verrebbero misurati non fra le effettive posizioni dell'effettore e del bersaglio, ma internamente fra posizioni attese in base ai modelli efferenti. Una strategia di questo tipo consentirebbe di iniziare il movimento con un programma motorio approssimativo, per poi rifinirlo grazie a cicli di aggiustamento interni che potrebbero iniziare ben prima che i segnali afferenti tornino indietro dai muscoli o dagli occhi. Si tratterebbe di un processo molto efficiente, perché consentirebbe di effettuare movimenti veloci in modo accurato, ma senza il carico cognitivo di «calcolare» esattamente tutti i parametri cinematici e dinamici del movimento fino alla sua fine.

Ma come è possibile raccogliere dati empirici a sostegno dell'utilizzo di questi inosservabili processi di simulazione interna? Forse i risultati più convincenti provengono proprio dallo studio dello spostamento della mano da parte di un'altra celebre paziente deafferentata, G.L. [Bard *et al.* 1999]. Il quadro clinico offerto da questa paziente è molto simile a quello di I.W.: una polineuropatia la ha sostanzialmente privata di sensibilità tattile e propriocezione al di sotto del collo, mentre ha preservato la sensibilità al dolore, al caldo e al freddo. Come nel caso di I.W., la paziente ha appreso nuovamente a usare le mani abbastanza bene, anche se non è mai stata in grado di imparare di nuovo a camminare. Va detto peraltro che la necessità di ricorrere alla sedia a rotelle è tipica di questo tipo di pazienti, per la maggior parte dei quali lo sforzo di controllare deambula-

fig. 3.8. A sinistra (A), le fasi di una prova nel paradigma *double-step*. Posto il dito sulla posizione iniziale (croce); compare un bersaglio da raggiungere (disco); durante lo spostamento della mano, il bersaglio si sposta; il partecipante deve modificare il movimento già iniziato per raggiungere la nuova posizione. La mano è visibile prima di iniziare il movimento ma diventa invisibile durante il movimento. A destra (B), i risultati medi relativi a un gruppo di partecipanti senza patologie e alla paziente G.L. Le prove in cui il bersaglio è a 12, 18 o 24 gradi sono prove normali. Quelle indicate con 18-12 e 18-24 si riferiscono a situazioni *double-step* in cui il bersaglio rispettivamente si sposta all'indietro, o in avanti come nell'esempio a sinistra.

fonte: Adattata da Bard *et al.* [1999].

zione e postura senza afferenze somatosensoriali raggiunge un livello non compatibile con una vita normale (in questo il caso di I.W. è ancora più eccezionale, dunque).

Naturalmente il controllo dei movimenti manuali da parte di G.L. è basato sulla visione. In un esperimento particolarmente ingegnoso, Bard e colleghi [*ibidem*] hanno studiato lo spostamento della mano verso un bersaglio da parte di G.L., usando un particolare paradigma che viene chiamato **paradigma del doppio passo** (*double-step*). Si tratta di una variazione al semplice compito di spostamento della mano in condizioni a circuito aperto (fig. 3.8A). Una volta posizionato il dito sulla posizione iniziale, il partecipante vede comparire il bersaglio e inizia a spostare la mano per raggiungerlo. Nel momento in cui inizia il movimento, un sistema ottico impedisce di vedere la mano mentre il bersaglio resta visibile. In alcune prove (le prove *double-step*), inaspettatamente, dopo l'inizio del movimento succede anche un'altra cosa: il bersaglio si sposta di alcuni centimetri, in avanti o indietro. Il compito del partecipante è correggere la traiettoria per arrivare alla nuova posizione del bersaglio.

I partecipanti senza patologie neurologiche riescono agevolmente a operare queste correzioni, a farlo in modo rapido e mentre stanno ancora muovendosi. La rapidità di queste correzioni e il fatto che avvengano anche se la mano non è visibile sono in accordo con una forma di controllo basata su modelli efferenti, che verrebbero modificati nel momento in cui si verifica un *double-step*. Nel caso dei partecipanti senza patologie non si può però escludere una spiegazione alternativa. Questi partecipanti potrebbero controllare l'azione usando un confronto multisensoriale fra due afferenze, vale a dire, la posizione visiva del bersaglio e la posizione della mano percepita per via propriocettiva-tattile. Questa possibilità non è molto plausibile se il movimento è molto veloce, ma rimane il fatto che il dato non è incontrovertibile. Questo dubbio ha motivato il lavoro di Bard e colleghi [*ibidem*], che sono stati in grado di dimostrare come G.L. riesca a modificare la traiettoria della mano nelle prove *double-step* in modo del tutto simile a quanto fanno i partecipanti senza patologie (fig. 3.8B). Ma se le cose stanno così, allora l'ipotesi del controllo afferente multisensoriale perde forza. Nel caso di G.L., infatti, la sensazione propriocettiva afferente sulla posizione della mano non c'è perché la paziente è deafferentata. Quindi qualsiasi modifica operata senza vedere la mano non può che basarsi su un modello interno efferente. Come dice Sherlock Holmes nel *Mastino dei Baskerville*, eliminato l'impossibile quello che resta, per quanto sembri improbabile, deve essere la verità.

6. CHI STA MUOVENDO IL MIO CORPO?

Nel saggio *Goodbye Kant*, il filosofo Maurizio Ferraris [2004] riporta uno degli aneddoti preferiti di uno dei grandi psicologi italiani della percezione, Paolo Bozzi. L'aneddoto riguarda un altro grande psicologo della percezione del secolo passato, Wolfgang Metzger. Durante la seconda guerra mondiale, distaccato a Cassino da militare nell'esercito tedesco, capitò a Metzger di usare la latrina del baraccamento dove alloggiava. Conclusa l'operazione, come è giusto tirò lo sciacquone. Proprio in quell'istante un proiettile d'artiglieria colpì le postazioni tedesche facendo esplodere il baraccamento – e Metzger ebbe l'impressione di essere stato lui, tirando la fatidica cordicella, a *causare* l'esplosione. Fenomeni come questo non sono rari né difficili da riprodurre. Ad esempio, uno di noi ha l'abitudine di usare una piccola bacchetta di legno per indicare, sulle diapositive presentate agli studenti, gli eventuali dettagli su cui di volta in volta

concentrare l'attenzione. Durante una di queste lezioni, proprio nell'istante in cui la bacchetta ha toccato lo schermo su cui era proiettata la diapositiva, nell'edificio è mancata la luce, il proiettore si è spento e la diapositiva è scomparsa – strappando una risata di sorpresa al docente per la fortissima impressione di avere *causato* la scomparsa della diapositiva con il suo gesto. Oppure, provate a fare questo semplice esperimento [Spizzo 1984]. Trovate un qualunque aggeggio che in certe condizioni faccia lampeggiare una lucina. Un metronomo digitale va benissimo, ma anche una di quelle lavastoviglie che avvisano quando il ciclo è finito con il lampeggiare di un led rosso. Osservando la luce, sforzatevi di sincronizzare un qualsiasi gesto con il ritmo di accensione e spegnimento. Ad esempio provate a schioccare le dita. Fate un po' di prove e noterete che in alcuni casi avrete l'impressione di essere voi a causare l'accensione della luce.

Che fenomeni di questo tipo abbiano attirato l'attenzione di un filosofo è ben comprensibile. Il dibattito sulla **causalità** percorre tutta la storia della filosofia da Aristotele, al dibattito fra Hume e Kant, fino all'epistemologia del Novecento. Storicamente, questo dibattito ha visto contrapposte due posizioni fondamentali [vedi ad esempio Abbagnano 1971]. Per la prima di queste, la causalità è una **connessione razionale**, tale per cui l'effetto è deducibile logicamente dalla causa. Per la seconda, invece, la causalità è una **connessione empirica**, riducibile a un rapporto osservabile di covariazione statistica fra causa ed effetto e quindi di prevedibilità. Per entrambe le posizioni lo sciacquone di Metzger rappresenta un bel grattacapo. Provate un po' a dedurre logicamente la bomba dalla latrina, o a sostenere l'esistenza di una covariazione osservabile fra sciacquoni ed esplosioni di artiglieria. Nello studio dei rapporti fra percezione, azione e cognizione i fenomeni come quelli descritti sopra sono invece interessanti perché dimostrano come i rapporti fra elaborazione cognitiva di un'intenzione, azioni, e conseguenze osservabili di queste azioni non siano così ovvi come potrebbe sembrare. Gli studiosi che si occupano di questi rapporti usano il termine **«agency»** per far riferimento al vissuto di essere l'origine della decisione di compiere un'azione e dunque di causare tale azione e le sue conseguenze. In quest'ultimo paragrafo del capitolo prenderemo in esame alcuni studi che hanno cercato di chiarire quali siano i processi percettivi e motori alla base della percezione di *agency*. Si tratta di argomenti assai rilevanti per gli obiettivi di questo libro perché, come vedremo, i processi in questione sono di natura squisitamente multisensoriale e coinvolgono direttamente rappresentazioni efferenti di tipo predittivo.

La ricerca che ha forse il merito maggiore nell'avere portato all'attenzione delle neuroscienze cognitive il problema della *agency* è un originalissimo

lavoro dovuto a un fisiologo dell'università della California, Benjamin Libet (1916-2007). Nel paradigma di Libet [Libet *et al.* 1983], i partecipanti devono svolgere un compito molto particolare: sollevare la mano posata su un tavolo nel momento in cui sentono il desiderio di farlo. Insomma, i partecipanti debbono semplicemente formulare l'intenzione di compiere un gesto e compierlo. Ma mentre fanno questo vengono raccolti dei dati, e questi dati sono di due tipi. Il primo tipo consiste di dati fisiologici. Sul cuoio capelluto dei partecipanti vengono applicati degli elettrodi per l'EEG, grazie ai quali viene registrata la modulazione nei potenziali elettrici generati dall'attività cerebrale (ne abbiamo accennato nel capitolo introduttivo al libro).

Contemporaneamente, sul braccio che esegue il compito vengono applicati altri elettrodi, grazie ai quali è possibile registrare i potenziali generati dall'attività muscolare (**elettromiografia**, EMG). Questi dati consentono quindi di avere informazioni sulla tempistica dei processi corticali associati all'azione intenzionale e dei processi muscolari che gestiscono obiettivamente il movimento. Il secondo tipo di dati raccolti nel paradigma di Libet consiste invece di una sorta di resoconto introspettivo, costituito da due giudizi: il **giudizio W** (da *Will*, volontà), relativo al momento in cui compare nella propria coscienza la consapevolezza di avere formulato la decisione di effettuare il movimento, e il **giudizio M** (*Movement*, movimento), relativo al momento in cui compare la consapevolezza che il movimento è effettivamente iniziato. Per quantificare questi due giudizi, Libet utilizzava uno stratagemma che è forse l'aspetto più originale di questo pionieristico studio. Accanto ai partecipanti poneva una specie di orologio con una lancetta che compiva una rivoluzione completa in circa tre secondi. I partecipanti dovevano prestare attenzione alla lancetta e, in alcune prove, riportarne la posizione nel momento in cui diventavano consapevoli dell'intenzione di muoversi (W); in altre prove invece dovevano riportarne la posizione al momento dell'inizio vero e proprio del movimento (M). Dunque nel paradigma di Libet vengono raccolti assieme dati sul decorso temporale di eventi fisiologici e di eventi relativi alla comparsa nella coscienza dei percetti corrispondenti.

I risultati dell'esperimento di Libet hanno diversi risvolti inaspettati. Innanzi tutto, esaminando l'EEG Libet notava che, diverse centinaia di millisecondi prima della contrazione del muscolo, l'attività corticale mostrava la tipica modulazione che precede il movimento, il cosiddetto **potenziale di preparazione** [vedi ad esempio Shibasaki e Hallett 2006]. Inoltre il tempo medio in cui si verificavano i giudizi W si collocava anche prima della contrazione muscolare. Questi due aspetti dei risultati sono in accordo con l'aspettativa che tutti abbiamo riguardo alla sequenza di eventi che dovrebbe caratterizzare il movimento

intenzionale: il movimento fisico inizia dopo che il soggetto ha percepito di avere l'intenzione di muoversi e dopo che il cervello ha generato l'attività necessaria per preparare il movimento. Ma non tutti i risultati di Libet sono ugualmente prevedibili. Infatti, confrontando il momento in cui iniziava la modulazione del potenziale di preparazione con il tempo medio dei giudizi W, Libet si accorse di una cosa assai meno ovvia. La modificazione dell'attività elettrica associata alla preparazione del movimento *precedeva* la consapevolezza dell'intenzione di muoversi, e non di poco – di circa 400 ms. Questo risultato è esattamente il contrario dell'aspettativa ingenua: uno si aspetterebbe che il Sé (la mente, la *res cogitans* di Cartesio) prenda la decisione di muoversi e che solo dopo il cervello (il corpo, la *res extensa* di Cartesio) esegua la decisione mandando il comando ai muscoli. Nel paradigma di Libet invece prima inizia l'attività cerebrale e solo dopo arriva il Sé cosciente, come se fosse la conseguenza e non la causa dell'attività cerebrale.

Questo aspetto dei risultati di Libet è dunque controintuitivo e in contrasto con la concezione normale della causalità e delle intenzioni. Come ha indicato lo stesso Libet, questo aspetto dei suoi risultati sembra andare contro alla nozione stessa di **libero arbitrio**. Ma è vera anche un'altra cosa. Se l'attività cerebrale precede la coscienza dell'intenzione, quest'ultima arriva comunque prima del movimento vero e proprio. Questo suggerisce la possibilità che la percezione di avere autonomamente deciso di muoversi sia il correlato fenomenico non della preparazione corticale del movimento, ma di un processo di ulteriore modifica del processo cerebrale. Infatti, in quei 200 millisecondi circa che separano la coscienza dell'intenzione dalla sua esecuzione c'è ancora tempo per decidere *come* fare il movimento (ad esempio, con la mano destra o sinistra) e per inibire i processi motori non selezionati. Dunque la scoperta di Libet potrebbe indicare che quello che percepiamo come libero arbitrio è, in effetti, una sorta di **libero veto** [Brass e Haggard 2007; vedi anche Obhi e Haggard 2004]. A sostegno di questa idea, studi successivi hanno dimostrato che la tempistica dei giudizi W non è ben correlata con quella del potenziale di preparazione classico, mentre corrisponde bene a quella di un'altra modulazione caratteristica dell'EEG, il **potenziale di preparazione lateralizzato** [Haggard e Eimer 1999]. Si ritiene che questo secondo aspetto dell'EEG, meno precoce, rifletta proprio i processi di selezione dell'effettore con cui eseguire l'azione [Brasil-Neto *et al.* 1992; Ammon e Gandevia 1990]. Ma per i fini di questo libro i dati di Libet contengono un risultato per certi versi ancora più interessante. Se si esamina la traccia elettromiografica che consente di misurare il tempo di inizio del movimento fisico, e la si confronta con i tempi medi dei giudizi M, dai dati di Libet emerge un altro

risultato sorprendente. Il giudizio M *precede* di circa 100 millisecondi l'inizio effettivo del movimento. Vale a dire, la coscienza del movimento arriva prima, non dopo, l'inizio del movimento stesso.

Se ci pensate, anche questo è profondamente controintuitivo. Infatti, il dato sembrerebbe indicare che quando percepiamo la conseguenza della nostra decisione di muoverci (ossia vediamo, e sentiamo a livello muscolare, che il braccio ha iniziato a muoversi), in realtà staremmo percependo il futuro. Il braccio inizierà a muoversi davvero solo fra alcune decine di millisecondi, e quello che stiamo percependo è una specie di pre-visione. Ma che le cose stiano esattamente in questo modo è attualmente oggetto di grande dibattito. Per molti ricercatori, ad esempio, il problema sta nel metodo utilizzato da Libet per misurare i tempi dei giudizi W e M. Questo metodo produrrebbe errori sistematici che potrebbero spiegare almeno in parte questi risultati come anche quelli relativi ai giudizi W [Danquah, Farrell e O'Boyle 2007; Joordens, van Duijn e Spalek 2002]. A questo proposito una posizione molto ben argomentata si deve al filosofo Dan Dennett [2003]. Secondo Dennett, i risultati di Libet sono sorprendenti solo perché noi abbiamo un fortissimo pregiudizio a considerare il Sé come un sistema unico dotato di un luogo mentale, una sorta di «teatro» cartesiano, dove i risultati delle operazioni di elaborazione vengono esibiti allo scopo di diventare oggetto della coscienza. Ma questa concezione del Sé potrebbe essere sbagliata. Molte considerazioni inducono a pensare a un **Sé non unitario** che emerge dall'interazione fra una varietà di processi in competizione e in continua evoluzione dinamica, ognuno in grado di generare rappresentazioni del proprio stato secondo la propria tempistica. Il gioco degli sfasamenti fra questi diversi processi potrebbe produrre anticipi e ritardi nel completamento di processi diversi. Ad esempio, abbiamo già visto (par. 5) come alcuni processi motori sembrino utilizzare processi efferenti che simulano il movimento e le sue conseguenze sensoriali in maniera predittiva, già prima che si renda disponibile il segnale afferente dai recettori. Questi processi predittivi potrebbero essere coinvolti nella percezione di quando inizia il movimento, interagendo con i processi visivi che leggono la posizione della lancetta o con altri processi ancora, in maniere che non capiamo del tutto.

La proposta di Dennett è interessante perché sposta l'interesse da un problema con implicazioni etico-filosofiche, il libero arbitrio, a un problema squisitamente empirico come è quello di valutare il ruolo del processo intenzionale nel funzionamento dei processi di integrazione multisensoriale e sensomotoria. Un fenomeno particolarmente rilevante, a questo proposito, è il cosiddetto raggruppamento o **unificazione intenzionale** (*intentional binding*) descritto dal

gruppo di Patrick Haggard a Londra [Haggard, Clark e Kalogeras 2002; Haggard e Clark 2003]. Questi ricercatori hanno usato il paradigma di Libet per misurare due tipi di resoconto introspettivo: un giudizio sul momento di inizio di un movimento, analogo al giudizio M di Libet, e un giudizio sul momento in cui si verifica la sua conseguenza sensoriale, nel caso di questi studi, un suono. In pratica, i partecipanti dovevano muovere un dito, questo produceva un suono, e immediatamente dopo, essi dovevano riportare la posizione della lancetta nel momento in cui erano diventati coscienti del primo o del secondo evento. Ma non è tutto qui. In alcune prove, il movimento era causato dall'intenzione del partecipante. Era, in un certo senso, «normale». In altre prove, invece, il movimento era provocato da uno stimolo magnetico applicato alla corteccia motoria, usando un apparecchio per la stimolazione magnetica transcranica (TMS; vedi par. 4 dell'introduzione). Dunque l'output motorio era confrontabile a quello del primo caso, meno un elemento cruciale: l'intenzione di muoversi. I risultati indicano che, nella condizione in cui il movimento è intenzionale, la percezione dell'inizio del movimento viene ritardata mentre la percezione del suono viene anticipata, rispetto agli analoghi giudizi sugli eventi isolati (sul tempo del solo movimento o del solo tono, non in sequenza). Quando il movimento è provocato dalla TMS, invece, avviene l'opposto: la percezione dell'inizio del movimento è anticipata mentre quella del suono viene ritardata. Dunque l'intenzione sembra modulare il processo di unificazione fra segnali sensoriali diversi, avvicinandoli nel tempo, mentre l'assenza di intenzione «marca» gli eventi come indipendenti e li allontana l'uno dall'altro nel tempo psicologico. Il fenomeno dell'*intentional binding* è stato replicato in numerosi studi [Engbert *et al.* 2007; Haggard e Cole 2007; Tsakiris e Haggard 2003] e si verifica indipendentemente dal canale sensoriale (visivo, uditivo, somatico) che veicola la conseguenza sensoriale dell'azione [Engbert, Wohlschlager e Haggard 2008].

Questi risultati suggeriscono che la percezione di *agency* non è una mera inferenza *a posteriori*, basata sulla co-occorrenza delle percezioni relative al proprio movimento e alla sua conseguenza nel mondo. Richiede invece che alla percezione del movimento sia effettivamente collegato un comando efferente corrispondente a un'intenzione. Tuttavia, fenomeni come lo sciacquone di Metzger suggeriscono che questa rappresentazione interna potrebbe essere specifica per l'intenzionalità del movimento, non per la causa che si intende provocare. Come dire, quello che conta è che intendevo muovere il braccio, non che volevo produrre un certo suono. Se le cose stanno così, dovrebbe essere possibile dimostrare sperimentalmente che la percezione di *agency* non viene distrutta quando l'esito dell'azione non corrisponde a quello atteso. Questa ipotesi è stata esaminata in un'articolata ricerca da Sato e Yasuda [2005]. Nel primo esperimento di questi autori, i partecipanti imparavano, durante una fase preliminare, che a due azioni arbitrarie (premere un tasto a sinistra o un tasto a destra) corrispondevano due conseguenze (un suono alto o un suono basso). Successivamente ai partecipanti venivano somministrate le prove sperimentali in cui una delle due azioni doveva venire eseguita e a essa seguiva il suono congruente con quello appreso o l'altro incongruente. Inoltre venivano manipolati anche il tempo fra la pressione del tasto e la presentazione del suono (il ritardo nella comparsa della conseguenza) e la modalità di esecuzione della pressione, autogenerata come negli esperimenti di Libet o in risposta a un segnale esterno. Dopo ogni prova sperimentale i partecipanti dovevano dare una valutazione introspettiva su quanto forte fosse la percezione di avere causato il suono, utilizzando una scala da 0 a 100 (dove 100 corrispondeva alla completa impressione di *agency*).

I risultati del primo esperimento di Sato e Yasuda indicano che l'impressione di *agency* si riduce da quasi 100 a circa 75 se il suono è incongruente rispetto a quando è congruente con quello appreso. Questo indica che la percezione di *agency* non è del tutto neutra rispetto a una specifica **aspettativa sulle conseguenze sensoriali** dell'azione. Tuttavia, rimane comunque cospicua anche quando l'esito dell'azione non è quello atteso. Molto forte è invece l'effetto del ritardo fra azione e suono: la percezione di *agency* scende da quasi 100 a circa 20 quando il ritardo cresce da 0 a 600 millisecondi. Da questo punto di vista, dunque, i risultati sono coerenti con fenomeni come lo sciacquone di Metzger. Se lo scoppio è ben sincronizzato con l'attivazione dello sciacquone, uno dovrebbe aspettarsi una forte percezione di *agency* anche se lo scoppio non è l'esito appreso. In aggiunta, Sato e Yasuda riportano anche che i dati non cambiano quasi per nulla se l'azione di premere i tasti viene effettuata in risposta a un segnale invece che autogenerata. Se vi siete procurati un metronomo e avete provato a replicare le vecchie osservazioni di Spizzo [1984], citate all'inizio di questo paragrafo, anche questo risultato del primo esperimento di Sato e Yasuda non dovrebbe stupirvi.

I due ricercatori giapponesi non si sono tuttavia fermati qui. In un secondo esperimento, hanno ripetuto le loro osservazioni utilizzando un compito un po' diverso, in cui i partecipanti dovevano premere i tasti in risposta a una configurazione che li induceva a commettere spesso degli errori, premendo il tasto che non intendevano premere. In questo caso, quindi, non veniva variata solo la congruenza con le aspettative sulla conseguenza dell'azione, ma anche l'esito motorio dell'intenzione. Sorprendentemente, i ricercatori riportano che anche

in questo caso la percezione di *agency* viene spostata solo di poco, da circa 100 nel caso dell'esito motorio corretto a circa 80 nel caso dell'esito erroneo. Quando invece la prova consiste di un esito motorio erroneo *unito* a una conseguenza incongruente, allora la percezione di *agency* si riduce drasticamente a circa 25. Questo risultato indica dunque che il processo coinvolto nell'*agency* non è strettamente specifico né per l'afferenza motoria relativa all'azione oggetto dell'intenzione, né per la sua conseguenza nel mondo, mentre è altamente sensibile all'interazione fra questi due fattori. In un lavoro recente, Moore e Haggard [2008] hanno concettualizzato la relazione fra *agency* e queste due componenti dell'azione intenzionale proponendo di distinguere fra un **processo predittivo** e un **processo inferenziale**. Il primo di questi corrisponde ai processi motori efferenti che predicono le conseguenze sensoriali del movimento, di cui abbiamo discusso nei paragrafi precedenti. Il secondo consisterebbe invece in un processo che valuta gli effetti delle proprie azioni sulla base dei segnali afferenti multisensoriali, confrontandoli con le intenzioni all'origine dell'azione effettuata. Secondo questa distinzione, il processo predittivo è intrinseco all'agente e contribuisce alla percezione di *agency* già prima dell'inizio dell'azione vera e propria; il processo inferenziale utilizza invece meccanismi cognitivi estrinseci all'agente che contribuiscono alla percezione di *agency* dopo che l'azione è stata effettuata.

Per studiare il peso relativo dei due processi sulla percezione finale di *agency*, Moore e Haggard hanno utilizzato il paradigma dell'*intentional binding*, manipolando sistematicamente la relazione fra un'azione e i suoi effetti. I risultati indicano che l'effetto di *intentional binding* viene modulato da questa relazione. In particolare, questi ricercatori hanno osservato che l'*intentional binding* si verificava anche nelle prove in cui l'azione non era seguita da una specifica conseguenza, ma solo quando questo effetto era altamente probabile in base alla statistica dell'associazione fra azione ed effetto stabilita durante le prove iniziali. Quando invece l'effetto non era probabile, il *binding* non si verificava. Assumendo che l'*intentional binding* sia un indicatore della percezione di *agency*, questi risultati suggeriscono che il processo predittivo è sufficiente per tale percezione, ma solo quando la previsione delle conseguenze dell'azione viene rinforzata da una covariazione sistematica fra azione e conseguenze che supporta il processo inferenziale. I risultati indicano quindi che i due processi interagiscono nel produrre un percetto integrato di *agency*.

PARTE SECONDA

Gli oggetti

CAPITOLO 4

L'interazione multisensoriale nella percezione di oggetti

La formazione delle unità percettive è un precursore fondamentale della percezione degli oggetti. I principi di organizzazione percettiva sottostanti a questo processo, benché solitamente descritti separatamente per i diversi sistemi sensoriali, sono certamente anche di natura multisensoriale. I meccanismi dell'integrazione multisensoriale che guida la fusione fra le molteplici informazioni riferite al medesimo oggetto sembrano essere descritti al meglio da modelli formali bayesiani.

1. L'ORGANIZZAZIONE PERCETTIVA

Nella vita di tutti i giorni, quando usiamo la parola «oggetto», ci sembra di sapere esattamente a cosa ci stiamo riferendo. Tuttavia definire cosa costituisca un oggetto non è un problema banale. Su Wikipedia, ad esempio, una voce filosofica definisce un oggetto come «una cosa, un'entità o un essere», e offre una lista di esempi che include «le piramidi, *Alfa Centauri*, il numero sette, la mia credenza che la predestinazione non esista, e la paura dei cani di tua madre».

Un oggetto, in questo senso, è qualsiasi cosa di cui si possa parlare o possa essere pensata. L'utilizzo comune invece tende a restringere il significato della parola a entità solide e inanimate, escludendo quindi gli animali, le entità immateriali e le rappresentazioni di oggetti in quadri e fotografie. Nell'ambito dello studio della percezione, tuttavia, il problema non è tanto definire cosa sia un oggetto e cosa non lo sia, quanto specificare attraverso quali processi e fasi di elaborazione il sistema percettivo formi rappresentazioni di entità che possono

diventare appunto oggetto di percezione [Peterson 2001], ossia venire «incontrate» [Metzger 1971] entrando a far parte della nostra esperienza cosciente con le caratteristiche fenomeniche dell'effettivamente esistente, e non meramente immaginato.

In questo capitolo ci occuperemo di un primo aspetto fondamentale della percezione di oggetti, ovvero del modo in cui avviene la formazione di unità organizzate (**organizzazione percettiva** o *grouping*) a partire dagli elementi disponibili nell'informazione in ingresso. Una conseguenza importante del processo di organizzazione è la costituzione di «figure», ossia di unità che si differenziano dalle altre unità e dalle non figure, ossia dalle zone della scena che rimangono prive di una forma specifica (sfondo). Il processo di organizzazione ha un ruolo indispensabile, perché gli elementi presenti nell'informazione non specificano in maniera univoca le caratteristiche percepite degli oggetti, e i sistemi percettivi debbono, in un certo senso, completare informazioni incomplete o rendere esplicite relazioni che non sono definite dai segnali in ingresso. L'informazione che dall'ambiente esterno arriva ai recettori sensoriali è caratterizzata da omogeneità e disomogeneità. Le disomogeneità identificano i **contorni** mentre le zone che non presentano variazioni vengono dette **regioni** (o *regioni uniformemente connesse*; Palmer e Rock [1994]). I contorni e le regioni costituiscono gli elementi fondamentali rilevati dai processi percettivi di base, tradizionalmente studiati soprattutto in riferimento al sistema visivo [Hubel e Wiesel 1968] (per una discussione specifica sui diversi modelli per l'estrazione di contorni visivi, vedi Bruce, Green e Georgeson [2003]).

Il processo di organizzazione percettiva si riferisce alle operazioni attraverso le quali i contorni e le regioni vengono raggruppati (*grouping*) in **unità percettive** (*unit formation*). L'organizzazione è necessaria perché gli elementi rilevati dai processi percettivi di base non identificano in modo univoco i margini degli oggetti nell'ambiente: molto spesso questi possono essere frammentari o incompleti. Considerate la figura 4.1, due immagini di oggetti (animati) i cui margini corrispondono a contorni frammentari a causa delle caratteristiche dello sfondo contro il quale sono posti. Nell'immagine a sinistra, un caso reale di mimetismo tratto da un documentario sulle tattiche di difesa dai predatori, c'è una sogliola adagiata sul fondo marino. In quella di destra, una variante di una celebre dimostrazione, ci sono un cane dalmata e uno sfondo costruito a bella posta per mimetizzarlo. In entrambi i casi, il processo di unificazione-segregazione può consentire, se ha successo, di localizzare il margine oggettuale completandolo nelle zone mancanti e attribuendogli una direzione rispetto alla distinzione figura-sfondo. Ad esempio, nel caso del cane dalmata, procedendo dall'alto verso

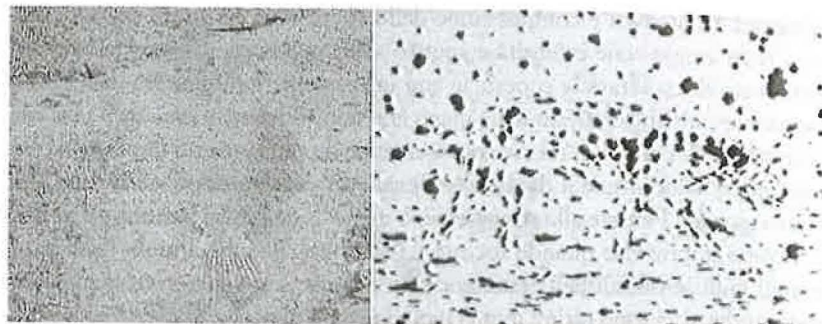



fig. 4.1. Due casi in cui i margini di un oggetto corrispondono a contorni frammentari nella struttura dell'informazione ottica.

il basso a un certo punto si verifica il passaggio dalla zona dell'immagine che costituisce lo sfondo a quella che costituisce la figura. Il margine appartiene alla figura (l'oggetto), mentre lo sfondo rimane privo di una forma specifica.

Nel contesto di questo libro, è importante sottolineare che il problema dell'organizzazione prospettiva può essere affrontato anche in una prospettiva multisensoriale. Considerate ad esempio la situazione rappresentata in figura 4.2, che richiama una delle infinite situazioni con cui potremmo avere a che fare nell'ambiente. La distribuzione della luce nell'immagine presenta cambiamenti

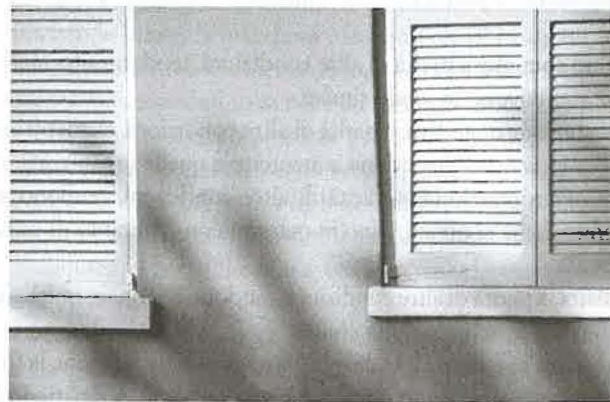



fig. 4.2. Omogeneità e disomogeneità nell'informazione che arriva ai recettori sensoriali. Notate i punti dell'immagine in cui l'intensità o la composizione spettrale della luce mutano bruscamente. Ora immaginate di esplorare il muro con un bastone. Durante il movimento, dove si verificheranno variazioni nelle forze applicate alla mano che regge il bastone? E nei suoni prodotti dal contatto fra il bastone e le superfici?

repentini di intensità e composizione dello spettro luminoso in alcuni punti, mentre in ampie zone intensità e spettro sono relativamente omogenei. Ora supponete di esplorare la superficie con un bastone. La dinamica delle forze (resistenze, attriti) registrata dalla mano che tiene il bastone sarà anch'essa caratterizzata da periodi di relativa omogeneità e da cambiamenti improvvisi. Ad esempio, quando urtate il davanzale. Uguale discorso vale per la sequenza dei suoni associati. Pensate allo sfregare più o meno costante del bastone sul muro, e al colpo improvviso quando toccate il davanzale. Benché il ruolo delle interazioni multisensoriali nell'organizzazione percettiva sia ancora relativamente poco studiato, ci sono ottimi motivi teorici per ritenere che la multisensorialità giochi un ruolo importante in questi processi percettivi. Soprattutto, ci sono svariati dati empirici a sostegno di questa idea. Prima di affrontare però il tema dell'organizzazione percettiva da una prospettiva multisensoriale è utile soffermarsi su quanto già sappiamo dell'organizzazione percettiva nella visione.

Il merito di avere per primi identificato alcuni dei principi che regolano i processi di unificazione-segregazione viene attribuito agli psicologi della scuola gestaltista, e in particolare a Max Wertheimer [1923]. Secondo Wertheimer, la formazione di unità percettive e la loro segregazione dallo sfondo obbedisce a sette principi o «fattori» di organizzazione, noti oggi come **principi di organizzazione di Wertheimer**. In estrema sintesi i principi sono:

1. **prossimità**: a parità di altre condizioni, tendono a formare unità gli elementi più vicini;
2. **somiglianza**: a parità di altre condizioni, tendono a formare unità gli elementi più simili;
3. **destino comune**: a parità di altre condizioni, tendono a formare unità gli elementi che si muovono in modo simile;
4. **impostazione oggettiva**: a parità di altre condizioni, elementi organizzati per formare una certa unità tendono a mantenere quella organizzazione;
5. **buona continuazione**: a parità di altre condizioni, tendono a formare unità gli elementi che si organizzano in margini con il minimo di cambiamenti di curvatura;
6. **chiusura**: a parità di altre condizioni, tendono a formare unità gli elementi che si organizzano formando unità chiuse;
7. **esperienza passata** o **abitudine**: a parità di altre condizioni, la formazione di una determinata unità percettiva può essere favorita dai ricordi o dalle conoscenze dell'osservatore.

A proposito dei primi sei fattori, Wertheimer insisteva nel sostenere la loro natura squisitamente percettiva, indipendente da conoscenze e memorie, e il

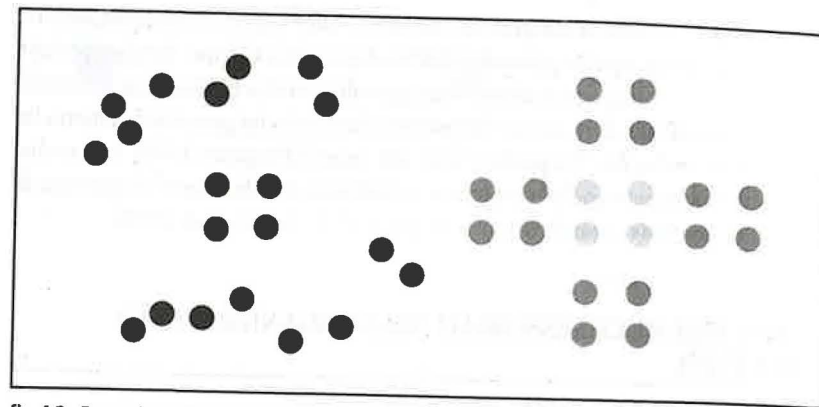



fig. 4.3. Formazione di unità per prossimità (sinistra) e somiglianza (destra).

loro funzionamento obbligatorio e automatico. Come dire: non si può fare a meno di organizzare il campo visivo in base alla prossimità o alla somiglianza (fig. 4.3), perché questi fattori di organizzazione riflettono un aspetto fondamentale, e condiviso da tutti i sistemi visivi, della percezione delle forme. Oppure: se la sogliola di figura 4.1 comincia a nuotare, non si può fare a meno di unificare (per destino comune) il movimento dei puntini sulla sua superficie. Ed ecco che automaticamente la silhouette del pesce si staglia dal fondo e diventa ben visibile. Il settimo fattore è invece in qualche maniera idiosincratico, perché dipende dall'esperienza individuale. Se ho passato molte ore a osservare la fauna marina con le pinne, la maschera e le bombole di ossigeno, potrei avere acquisito una particolare familiarità con la forma delle sogliole e con gli elementi meno facilmente mimetizzabili (osservate ad esempio dove sta la pinna caudale nella figura). In questo caso, l'esperienza passata potrebbe consentirmi di organizzare efficacemente elementi percettivi che, altrimenti, non formerebbero spontaneamente un'unità percettiva in base agli altri fattori. Ma senza questo particolare tipo di competenza, trovare il pesce rimane assai difficile. Non a caso, presumibilmente, fino a che non avete letto cosa dovevate cercare nell'immagine non ci avete visto proprio nulla.

In accordo con la posizione di Wertheimer, diversi lavori recenti suggeriscono che l'organizzazione in base a fattori come la somiglianza o la buona continuazione potrebbe essere implementata da meccanismi neurali presenti nelle aree visive primarie [Kapadia *et al.* 1995; Peterhans e von der Heydt 1989]. Plausibilmente, questi meccanismi sono coinvolti sia nell'interpolazione di margini figurali non completamente definiti da contorni nell'assetto ottico,

sia nel completamento di margini figurali dietro altre superfici (completamento amodale), un altro aspetto pervasivo dell'ecologia della visione. Sembra plausibile inoltre che un discorso analogo rimanga valido anche in modalità sensoriali diverse dalla visione. Può essere dimostrato che i principi gestaltisti valgono in larga misura anche per l'organizzazione dei suoni [Bregman 1990] e ricerche recenti hanno iniziato anche a esplorare somiglianze e differenze fra i principi di organizzazione visivi e tattili [Harrar e Harris 2007; Serino *et al.* 2008].

2. PROCESSI MULTISENSORIALI NELL'ORGANIZZAZIONE PERCETTIVA

Se da un lato sembra assodato che processi di unificazione-segregazione sono implementati dai meccanismi percettivi di base in diverse modalità di senso, allo stato attuale è meno chiaro se, e in che misura, l'organizzazione percettiva possa essere influenzata da vere e proprie interazioni multisensoriali. Da un punto di vista teorico, sembrerebbe logico attendersi che la multisensorialità abbia un ruolo importante, proprio perché informazione incompleta o ambigua in un canale unisensoriale potrebbe essere completa o meno ambigua in un altro. Considerate ancora l'esempio della figura 4.2. Un problema che il sistema visivo incontra nell'organizzare gli elementi disponibili è costituito dalle ombre. Le ombre sono dotate a tutti gli effetti di contorni ma tali contorni non vanno evidentemente attribuiti al margine di un oggetto segregato da uno sfondo, ma piuttosto a un salto nell'illuminazione. L'informazione tattile, acquisibile ad esempio esplorando col bastone la superficie del muro in corrispondenza dell'ombra, potrebbe allora giocare un ruolo importante perché la discontinuità presente nel segnale visivo non è presente nel segnale tattile, che potrebbe quindi fungere da fattore disambiguante. Almeno un lavoro suggerisce che il tatto può avere un ruolo simile a questo nella percezione dell'ombreggiatura di una superficie [Adams, Graf ed Ernst 2004]. Ne parliamo in dettaglio nel paragrafo 5 di questo capitolo.

O ancora, immaginate di osservare la scena nella figura 4.2 in condizioni di bassa illuminazione, ad esempio al tramonto. Se l'illuminazione è scarsa, è possibile che i margini di una superficie non corrispondano più a discontinuità nell'informazione, ossia a contorni. Ma naturalmente la scarsità di illuminazione non ha effetto sull'informazione tattile o acustica, e quindi dovremmo aspettarci che questi canali sensoriali possano contribuire a determinare la corretta organizzazione della scena. Diverse prove empiriche suggeriscono che questo tipo

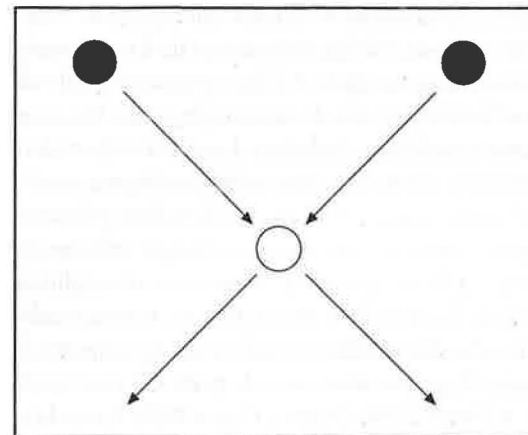



fig. 4.4. Traiettorie dei dischi stimolo nell'esperimento di Sekuler, Sekuler e Lau [1997]. Possono essere percepiti l'incrocio dei dischi, uno dei quali passa attraverso o sotto l'altro (scorrimento), oppure l'urto e il rimbalzo con cambio di direzione (rimbalzo). L'aggiunta di un suono aumenta moltissimo la probabilità della seconda soluzione.

di interazione multisensoriale è anch'esso possibile. Considerate, ad esempio, la situazione illustrata nella figura 4.4. Osservando un'animazione in cui due dischi descrivono le traiettorie illustrate nella figura è possibile percepire due eventi diversi: l'incrocio fra due oggetti che hanno traiettorie rettilinee, con uno degli oggetti che passa sotto o attraverso l'altro (scorrimento), oppure l'urto fra due oggetti che dopo essersi toccati rimbalzano indietro con una traiettoria a novanta gradi rispetto a quella originale (rimbalzo). Per il fattore della buona continuazione, le persone tendono a preferire la prima soluzione. Tuttavia l'ambiguità dello stimolo, dove non ci sono indizi che un disco stia davanti all'altro, unita alla nostra conoscenza che gli oggetti non possono compenetrarsi, non impedisce che anche la seconda soluzione venga percepita, anche se con minore frequenza. Un celebre esperimento di alcuni anni fa [Sekuler, Sekuler e Lau 1997] ha mostrato che la seconda opzione diventa dominante se allo stimolo viene aggiunta una componente multisensoriale. Basta aggiungere, nel momento dell'incrocio, un breve click che suggerisca un urto. La discontinuità nel canale uditivo sembra dunque contribuire a risolvere l'ambiguità presente in quello visivo.

Va sottolineato che anche questo effetto sembra essere genuinamente percettivo e non legato a semplice suggestione. Infatti, è stato dimostrato [Sanabria *et al.* 2004] che l'interazione multisensoriale si manifesta anche in una misura indiretta, che non richiede al partecipante di riferire esplicitamente la propria esperienza. I partecipanti di questo secondo studio dovevano giudicare, il più rapidamente possibile, da quale lato di una superficie occludente emergeva un

disco bianco che veniva visto precedentemente avvicinarsi alla superficie. Contemporaneamente, e in direzione opposta, veniva mostrato anche il movimento di un disco nero. Il disco bianco poteva emergere dal lato opposto a quello di avvicinamento (condizione scorrimento) oppure di nuovo dallo stesso lato, con movimento in direzione contraria (condizione rimbalzo). I risultati dello studio indicano chiaramente che le risposte sono più rapide nella condizione scorrimento, se non viene aggiunto un suono, mentre sono più rapide nella condizione rimbalzo se c'è il suono. L'organizzazione percettiva viene dunque influenzata dall'informazione multisensoriale anche se non viene chiesto in modo esplicito di descrivere la propria esperienza. Il ruolo delle interazioni fra visione e udito nell'organizzazione percettiva che sfocia nella percezione del movimento di un oggetto è attualmente oggetto di grande attenzione da parte dei ricercatori [Bushara *et al.* 2003; Freeman e Driver 2008; Grassi e Casco 2009; Kawachi e Gyoba 2006; Lyons *et al.* 2006; Sanabria *et al.* 2005].

Gli studi appena discussi indicano dunque che l'informazione in un canale unisensoriale può orientare la scelta di un'alternativa percettiva in condizioni di ambiguità in un altro canale, favorendo un'organizzazione degli elementi disponibili a scapito di altre [vedi anche Wijntjes *et al.* 2009]. Ma la **disambiguazione multisensoriale** è probabilmente solo un caso particolare delle interazioni multisensoriali che caratterizzano i processi di organizzazione. Un effetto particolarmente sorprendente, a questo proposito, è stato riportato già alcuni anni fa da Shams, Kamitani e Shimojo [2000]. I partecipanti a questo esperimento vedevano dei brevi flash di luce, presentati in sequenza, e dovevano giudicare quanti flash avevano visto. I flash potevano essere da 1 a 4. In alcune prove, oltre ai flash erano presentati anche dei suoni (*beep*). Anche i suoni potevano essere da 1 a 4. I ricercatori hanno trovato che i partecipanti erano molto accurati nello svolgere il compito quando i *beep* non erano presenti, mentre commettevano errori sistematici quando dovevano giudicare un singolo flash ma contemporaneamente venivano presentati *beep* in numero maggiore di uno. In particolare, quando con un flash solo venivano presentati due *beep* quasi tutti i partecipanti giudicavano di avere visto due flash. E quando ne venivano presentati tre un numero abbastanza cospicuo riteneva ancora di avere visto tre flash [Shams, Kamitani e Shimojo 2002]. La doppia o tripla discontinuità presente nel canale acustico dunque influenzava l'unificazione percettiva dell'oggetto, provocando, in alcuni casi, il suo sdoppiamento a dispetto dell'assenza di una doppia discontinuità visiva, del tutto non ambigua anche se forse poco affidabile a causa della presentazione molto rapida. A sostegno della generalità del principio, è stato riportato che l'effetto funziona anche quando, al posto del segnale sonoro, vengono somministrati uno o più stimoli tattili sotto forma di brevi pressioni su un polpastrello [Violentyev, Shimojo e Shams 2005].

Tra i fattori coinvolti nell'unificazione-segregazione, infine, vanno citate anche le caratteristiche dell'informazione sensoriale che sono indicative della stratificazione in profondità fra superfici. Ad esempio, quando un oggetto è collocato nell'ambiente davanti a una superficie di sfondo, sono tipicamente disponibili informazioni spaziali binoculari (parallasse binoculare) e monoculari (parallasse di movimento, configurazioni di contorni che indicano occlusione, ombreggiature) sulla stratificazione fra superfici e sulla struttura tridimensionale dell'oggetto che diventa figura. Come vedremo nella parte sulla percezione multisensoriale dello spazio, anche a questo livello sembra plausibile ipotizzare un coinvolgimento di specifiche interazioni multisensoriali.

3. FONDERE I SEGNALI MULTISENSORIALI

Molto spesso, quando abbiamo davanti un oggetto, la percezione è una sorta di «fusione» delle informazioni che provengono dai diversi canali sensoriali. Ad esempio, nel formarci una rappresentazione della grandezza di un oggetto possiamo tenere conto sia del segnale visivo sia di quello tattile. La storia della ricerca sulla percezione multisensoriale della grandezza degli oggetti è particolarmente istruttiva, perché alcune osservazioni degli anni '60 [Rock e Victor 1964] erano state originariamente interpretate come dimostrazioni di dominanza visiva e non di fusione: in presenza di conflitto tra la forma vista e quella toccata, i resoconti dei soggetti tendevano a coincidere con il segnale visivo (ne abbiamo già parlato nell'introduzione). Diverse ricerche successive, tuttavia, in condizioni solo di poco diverse hanno mostrato che la dominanza visiva non è sempre totale [Powert 1980], che si può verificare anche dominanza tattile [Hershberger e Misceo 1996], e che la dominanza dipende da vincoli attentivi e legati al compito [Heller *et al.* 1999].

Quest'ultimo risultato in particolare appariva in accordo con un'ipotesi diversa da quella della dominanza visiva, che veniva chiamata **ipotesi della modalità più appropriata** (*modality appropriateness*). Secondo questa ipotesi, il sistema percettivo sceglie di volta in volta, in base alla situazione e al compito, la modalità più appropriata su cui basarsi per la percezione di una determinata caratteristica oggettuale [Welch e Warren 1986]. La letteratura recente ha unificato questi punti di vista, confermando che la percezione multisensoriale di un'estensione vista o toccata è *sempre* il risultato di una fusione dei segnali

unisensoriali [Ernst e Banks 2002; Gepshtein e Banks 2003], ma che il peso relativo di questi segnali varia in funzione di altre proprietà dello stimolo, come ad esempio la sua affidabilità. Questa conclusione, che sembra valida in generale perlomeno negli adulti (mentre potrebbe non valere nei bambini, vedi Gori *et al.* [2008]), si è rivelata molto adatta a spiegare anche molte altre proprietà oggettuali che possono essere multisensoriali: ad esempio la forma tridimensionale [Helbig ed Ernst 2007] e l'inclinazione in profondità [Ernst *et al.* 2000].

Le proprietà di un oggetto non si limitano alle sue caratteristiche strettamente spaziali. Come abbiamo già visto sempre nell'introduzione al testo, persino l'impressione che una superficie sia liscia o ruvida non dipende solo, come è ovvio, dal tatto e dalla visione ma anche dall'udito [Guest *et al.* 2002; Jousmäki e Hari 1998]. Inoltre, quando solleviamo un oggetto, la grandezza dell'oggetto, specificata dal canale visivo, può contribuire a farcelo giudicare più o meno pesante [Kawai *et al.* 2007]. Il contributo della visione alla percezione del peso di un oggetto è un fenomeno noto da lungo tempo: nella cosiddetta **illusione grandezza-peso** [Charpentier 1891], di due oggetti di peso uguale, quello visivamente più piccolo sembra pesare di più. Ma l'effetto della grandezza sul peso si verifica anche quando la grandezza dell'oggetto viene registrata col tatto invece che con la vista. Ad esempio, un oggetto viene giudicato più leggero se è stato afferrato aprendo molto le dita rispetto a quando viene afferrato aprendole poco; se è stato afferrato con una presa di potenza (tutte e cinque le dita) rispetto a quando viene afferrato con una presa di precisione (solo due dita); e anche se viene afferrato in modo da avere una superficie di contatto ampia rispetto a quando viene afferrato con una superficie di contatto piccola [Flanagan e Bandomir 2000]. La percezione del peso di un oggetto può dunque anch'essa risultare da un'interazione complessa fra visione, sensazione di sforzo muscolare, tatto esplorativo.

4. L'APPROCCIO BAYESIANO

Una gran parte dei dati appena illustrati è ben spiegata da un modello formale derivato dall'**approccio bayesiano** alla percezione. Secondo tale approccio, che costituisce la formulazione contemporanea di idee già proposte nell'Ottocento da Helmholtz [1867] e nel secolo scorso da Rock [1985], il processo percettivo consiste nella scelta dell'interpretazione più probabile per l'informazione sensoriale disponibile in un dato momento. Tale scelta è determinata dal prodotto di due cose: l'informazione sensoriale disponibile, e ciò che si sa

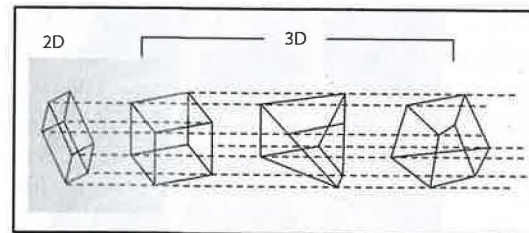



fig. 4.5. L'ambiguità di una proiezione bidimensionale rispetto alla struttura tridimensionale che l'ha generata. A ogni proiezione possono corrispondere infinite forme tridimensionali più o meno regolari.

già sull'ambiente e gli oggetti. Per la precisione, l'informazione sensoriale corrisponde a quella che, nel teorema di Bayes, viene chiamata la **verosimiglianza**, ossia, nel nostro caso, la probabilità che sia disponibile una certa informazione dato un certo stato di cose nel mondo; mentre le conoscenze già maturate costituiscono quella che viene chiamata la **probabilità a priori** (ossia precedente all'acquisizione di nuova informazione). Il prodotto di verosimiglianza e probabilità a priori definisce la **probabilità a posteriori**, nel nostro caso, la probabilità di un certo stato di cose nel mondo, data l'informazione disponibile. I valori di verosimiglianza e probabilità a priori per cui la probabilità a posteriori è massima definiscono quale stato di cose viene percepito.

Per comprendere appieno la logica dell'approccio bayesiano alla percezione è utile considerare innanzi tutto il **ruolo delle conoscenze a priori** nel processo percettivo. Un esempio classico [vedi Helbig ed Ernst 2008] è la percezione della struttura tridimensionale di un oggetto semplice a partire dalla sua proiezione retinica. Siccome la retina è bidimensionale, la proiezione è sempre ambigua rispetto alla struttura tridimensionale che l'ha generata. Consideriamo, ad esempio, delle strutture tridimensionali costruite con del filo di ferro come quelle in figura 4.5. Se consideriamo la relazione fra queste strutture e le loro potenziali proiezioni bidimensionali, ci possiamo rendere immediatamente conto che la stessa configurazione di contorni sulla retina potrebbe essere causata da un cubo o da infiniti solidi irregolari. In queste condizioni dunque non esiste una maniera univoca di derivare la struttura tridimensionale dell'oggetto a partire dalla sua proiezione sulla retina. La soluzione può essere solo probabilistica, e secondo l'approccio bayesiano percepire significa appunto scegliere la soluzione a cui è associata la probabilità (a posteriori) più elevata.

Per il teorema di Bayes, tale probabilità è proporzionale al prodotto della verosimiglianza per la probabilità a priori. Il risultato percettivo dipende dunque dalla probabilità di avere quella configurazione sulla retina quando sia

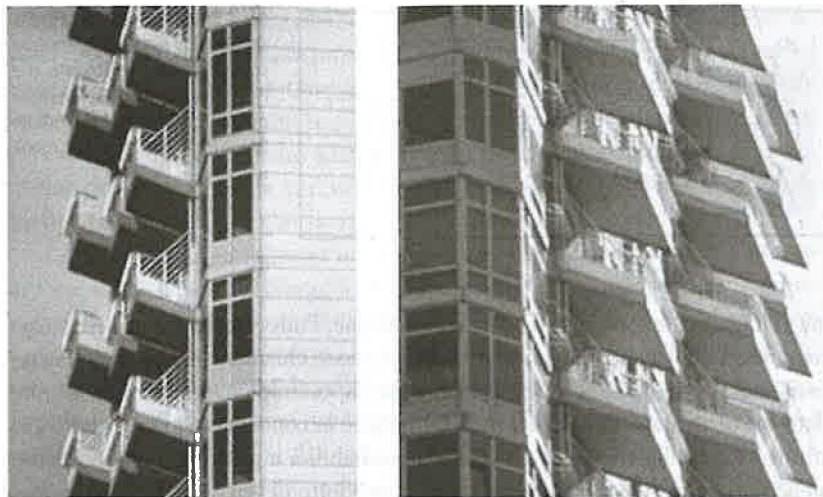



fig.4.6. Fotografie del condominio «The Future» a Manhattan. Le foto sono prese dal marciapiede sul lato opposto della strada rispetto all'edificio, da due posizioni rispettivamente a sinistra e a destra dell'edificio stesso.

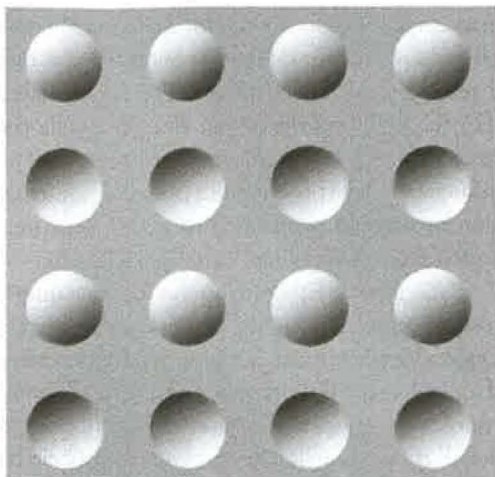
fonte: Halper [1997].

presente una certa struttura tridimensionale (funzione di verosimiglianza), e dalla probabilità a priori di quella stessa struttura. Nel caso esaminato, a causa dell'ambiguità proiettiva che caratterizza oggetti tridimensionali come questi, la distribuzione delle funzioni di verosimiglianza è uniforme (la probabilità della configurazione 2D è sempre la stessa per tutte le possibili strutture 3D). Nel determinare il percetto gioca dunque un ruolo cruciale la probabilità a priori. Non è difficile dimostrare, ad esempio, che il sistema percettivo umano preferisce, quando possibile, interpretare gli angoli formati da due contorni sulla retina come angoli retti di superfici oggettuali appropriatamente orientate nello spazio. Un esempio molto efficace di questa preferenza è visibile nella figura 4.6 [da Halper 1997], dove vengono confrontate due fotografie di un edificio con balconi a forma di parallelepipedo. L'edificio, che si trova a New York, merita una deviazione per un piccolo esperimento naturale. Osservando i balconi dal marciapiede di fronte, questi non appaiono come sono veramente, ma sembrano avere una normale pianta rettangolare (preferenza per l'angolo retto). Allo stesso tempo, questi balconi sembrano assai bizzarri: appaiono inclinati verso il basso o verso l'alto, a seconda che si guardi dalla destra o dalla sinistra dell'edificio. In termini bayesiani, il fenomeno può essere interpretato in questa

maniera: l'immagine è proiettivamente ambigua (un parallelepipedo sulla retina può essere causato da infinite superfici con diversa forma o inclinazione), ma il sistema «sa» che le superfici con angoli retti tendono a essere più probabili (la probabilità a priori è più alta per quella configurazione rispetto alle altre irregolari). A parità di informazione retinica, il percetto viene quindi fortemente influenzato dalla probabilità a priori, al punto da farci vedere una soluzione che, a un altro livello, sappiamo benissimo essere impossibile (per ovvi motivi non si costruiscono balconi inclinati verso il basso).

L'esempio è istruttivo, perché ci fa capire che l'approccio bayesiano non sostiene, come si potrebbe pensare, che si percepisce quello che si sa sul mondo. Questa versione rozza dei rapporti fra percezione e conoscenza è sicuramente sbagliata, come dimostrano anche tutti gli innumerevoli effetti percettivi in cui conoscenze e aspettative non ci impediscono di continuare a cadere vittima dell'illusione. Quello che sostiene l'approccio bayesiano è invece che il sistema ha incorporate delle aspettative molto generali sulle regolarità statistiche che caratterizzano gli oggetti e l'ambiente. Ad esempio, che gli oggetti tendono a essere compatti e regolari, ad avere dimensioni comprese in una certa gamma, e così via. A parità di informazioni sensoriali, queste conoscenze a priori possono orientare il processo percettivo e influenzarne l'esito. Ma solo a parità di informazioni sensoriali. Ponendosi esattamente davanti all'edificio e il più possibile vicino ai balconi più bassi, la forma percepita non è più quella di un balcone rettangolare ma quella, irregolare, del balcone reale a forma di parallelepipedo. Nell'approccio bayesiano, questo avviene perché avvicinandosi all'edificio diventano disponibili altre informazioni visive (funzioni di verosimiglianza), incompatibili con l'interpretazione illusoria. Al crescere della verosimiglianza associata all'interpretazione corretta della forma del balcone, si riduce dunque l'influenza dell'aspettativa a priori fino al punto in cui la soluzione più probabile diventa quella corretta.

Il ruolo della probabilità a priori è dunque, nell'approccio bayesiano, quello di contribuire a risolvere ambiguità più o meno parziali nell'informazione sensoriale. Quanto più quest'ultima è ricca e non ambigua, tanto meno la componente a priori influenza il risultato finale. Infatti, in condizioni naturali non abbiamo tipicamente difficoltà a percepire correttamente anche le forme irregolari, quando queste sono bene specificate dall'informazione sensoriale. E non solo. Dato che le probabilità a priori dipendono da conoscenze che vengono incorporate con l'esperienza, individuale o della specie, l'approccio bayesiano implica che proprio le informazioni sensoriali dovrebbero, a lungo andare, anche essere in grado di modificare le probabilità a priori.

fig. 4.7. Il sistema visivo tende a interpretare il chiaroscuro assumendo che l'illuminazione provenga dall'alto. I dischi nella prima e nella terza riga della figura sono identici a quelli nella seconda e nella quarta riga, a meno di una rotazione di 180 gradi. Tuttavia nelle righe uno e tre si vedono delle collinette, mentre nelle righe due e quattro degli avvallamenti. Secondo l'approccio bayesiano l'effetto è dovuto all'alta probabilità «a priori» che la luce provenga dal sole (ossia appunto dall'alto).

Una ricerca molto elegante, a questo proposito, è stata condotta da Wendy Adams dell'Università inglese di Southampton [Adams, Graf ed Ernst 2004]. La Adams ha sfruttato l'effetto rappresentato in figura 4.7, in cui l'interpretazione di una superficie come concava o convessa sembra dipendere dalla cosiddetta **assunzione della luce dall'alto**. Nell'approccio bayesiano, questa assunzione corrisponde all'elevata probabilità a priori che la luce provenga dal sole, la quale influenza la conclusione sulla forma tridimensionale in condizioni in cui la convessità o concavità deve essere inferita probabilisticamente a partire dal solo chiaroscuro (*shading*) nell'immagine. Nel suo esperimento, la Adams ha misurato, per ogni partecipante, la preferenza per l'interpretazione concava o convessa in funzione dell'orientazione del chiaroscuro più o meno verso l'alto o il basso. A questa misura iniziale ha poi fatto seguire una fase di addestramento multisensoriale, in cui i partecipanti toccavano con la mano superfici effettivamente concave o convesse, associate a configurazioni visive ruotate rispetto a quelle iniziali (e quindi non più coerenti con l'assunzione che la luce provenga dall'alto). Infine la Adams ha misurato nuovamente le preferenze per l'interpretazione concava o convessa in funzione del chiaroscuro. I risultati hanno evidenziato che le preferenze per l'una o l'altra interpretazione non erano più completamente coerenti con l'assunzione che la luce provenga dall'alto, ma si erano spostate nella direzione dell'interpretazione imparata durante l'addestramento visuotattile. Il risultato è quindi in accordo con l'idea che l'esperienza di nuove associazioni fra tatto e vista può modificare l'aspettativa a priori dell'osservatore, alterando in queste condizioni l'interpretazione del chiaroscuro. In termini bayesiani, l'esperienza multisensoriale avrebbe modificato la distribuzione delle probabilità a priori relative alla direzione da cui proviene la luce, e di conseguenza anche quella a posteriori che orienta la scelta della soluzione percepita.

5. I PESI DEI SENSI

L'approccio bayesiano può essere utilizzato per derivare un modello in cui la fusione multisensoriale consente di utilizzare al meglio le informazioni unisensoriali disponibili. Questo modello viene definito **modello di integrazione ottimale**. La derivazione formale del modello va al di là degli obiettivi del libro [per maggiori dettagli, vedi Helbig ed Ernst 2008]. Ci limiteremo quindi a richiamare le sue caratteristiche fondamentali. Assumendo distribuzioni uniformi per le probabilità a priori, è possibile dimostrare che l'interpretazione ottimale dei segnali multisensoriali corrisponde alla somma dei segnali unisensoriali pesati in base alla loro affidabilità relativa. Ad esempio, se consideriamo il caso della percezione della struttura tridimensionale di un oggetto in presenza di segnali visivi e tattili, il percetto multisensoriale (ossia l'interpretazione integrata per cui la probabilità a posteriori è massima) sarà

$$S_{vt} = W_v S_v + W_t S_t$$

dove S_v e S_t sono le migliori interpretazioni che il sistema avrebbe fatto, sulla base dei soli segnali visivi e tattili, S_{vt} è l'interpretazione integrata, e W_v, W_t sono i pesi relativi rispettivamente assegnati ai due segnali unisensoriali (dato che si tratta di pesi relativi, $W_v + W_t = 1$ e l'operazione è più precisamente una media ponderata). I pesi sono proporzionali all'affidabilità dei segnali unisensoriali, ossia sono tanto più grandi quanto più piccola è la variabilità delle funzioni di verosimiglianza associate al segnale unisensoriale. Pertanto tutti i segnali unisensoriali vengono utilizzati, ma viene dato più peso a quelli maggiormente affidabili. Inoltre, per le proprietà statistiche del modello, la precisione della stima multisensoriale integrata è necessariamente maggiore delle precisioni associate alle singole stime unisensoriali. L'integrazione dei diversi canali sensoriali è dunque ottimale in quanto: *a*) utilizza al meglio l'informazione fornita da tutti i canali sensoriali disponibili; e *b*) il percetto integrato è meno soggetto a variabilità casuale rispetto a quanto non siano i singoli segnali sensoriali prima dell'integrazione.

Il modello dell'integrazione ottimale può essere utilizzato per derivare previsioni quantitative sui risultati della fusione di più canali sensoriali. Ad esempio, è possibile prevedere la grandezza percepita di un oggetto quando questa sia specificata simultaneamente dalla vista e dal tatto, nonché la precisione associata a tale percetto. Nell'ultimo decennio, alcune ricerche svolte nel laboratorio di Marty Banks all'Università della California hanno dimostrato che queste previsioni sono in buon accordo con le misure ottenute nei compiti che richiedono fusione multisensoriale. In una di queste ricerche, Ernst e Banks [2002] hanno chiesto a dei volontari di decidere quale di due oggetti presentati in sequenza fosse più grande (per la precisione, quale di due superfici fosse più in rilievo). Il compito veniva svolto o solo visivamente, o solo col tatto (attivo), oppure in condizioni multisensoriali. Lo stimolo visivo era generato utilizzando una tecnica (stereogramma di punti casuali) che consente di variare l'affidabilità dell'informazione visiva sulla grandezza (in sostanza, veniva manipolata la percentuale di punti coerente con un oggetto di una certa grandezza, aumentando o diminuendo il «rumore» associato al segnale visivo). Lo stimolo tattile era invece generato utilizzando un'apparecchiatura in grado di applicare forze appropriate sulle dita del partecipante. I due canali sensoriali erano dunque controllati in modo indipendente, variando l'affidabilità del segnale visivo mentre rimaneva costante quella del segnale tattile.

Per stimare la precisione delle stime unisensoriali di grandezza, Ernst e Banks hanno calcolato, sulla base delle discriminazioni di grandezza unisensoriali, le soglie differenziali relative ai due compiti. Ad esempio, nel caso del segnale visivo, tanto maggiore la percentuale di rumore nello stimolo, tanto più alta la soglia differenziale (differenza di grandezza necessaria perché si percepisca il cambiamento) e dunque tanto minore la precisione della stima. A partire dalle precisioni unisensoriali così calcolate, i due ricercatori hanno quindi derivato previsioni sulla grandezza percepita in condizioni multisensoriali e sulla precisione a essa associata. Naturalmente, queste previsioni sono state ricavate separatamente per i diversi livelli di rumore visivo. I risultati hanno indicato un eccellente accordo con le previsioni del modello dell'integrazione ottimale. Quando il segnale visivo era altamente preciso (assenza di rumore, dunque previsione di un peso visivo elevato), il percetto multisensoriale era dominato dalla visione e dipendeva solo in minima parte dal segnale tattile. Quando il segnale visivo era meno preciso, invece, il peso relativo del segnale tattile aumentava, fino a che il tatto non diventava a sua volta predominante quando il segnale visivo era diventato molto impreciso. Ernst e Banks hanno dunque dimostrato che, nella percezione multisensoriale della grandezza, è possibile ottenere dominanza visiva ma anche dominanza tattile, in funzione della precisione associata ai due segnali unisensoriali. Oltre a confermare le previsioni del modello ottimale, questo spiega dunque i risultati contraddittori che abbiamo discusso all'inizio di questo paragrafo [Rock e Victor 1964; Power 1980; Hershberger e Misceo 1996; Heller *et al.* 1999].

Infine, un'ulteriore conferma del modello ottimale è venuta, nel lavoro di Ernst e Banks, dal confronto fra le soglie differenziali nella condizione multisensoriale e nelle condizioni unisensoriali. Come abbiamo accennato sopra, queste soglie sono inversamente proporzionali alla precisione delle stime percettive associate: tanto più bassa la soglia, tanto maggiore la precisione. Una previsione specifica del modello ottimale è che la precisione della stima multisensoriale dovrebbe essere migliore delle precisioni unisensoriali. Anche in questo caso, la previsione è stata confermata: le soglie differenziali multisensoriali erano effettivamente minori rispetto a quelle unisensoriali.

In un altro esperimento dello stesso gruppo di ricerca [Gephstein e Banks 2003], lo stesso metodo è stato applicato a una situazione in cui l'affidabilità dello stimolo visivo non era manipolata inserendovi del rumore ma variandone invece la geometria rispetto al punto di vista. La logica di questa manipolazione sperimentale è illustrata nella figura 4.8. La stima visiva dell'estensione fra due superfici è molto affidabile quando questa è ortogonale alla linea di mira. In questo caso, infatti, l'estensione da stimare non viene distorta sulla proiezione retinica. Ruotando le due superfici rispetto alla linea di mira, la proiezione

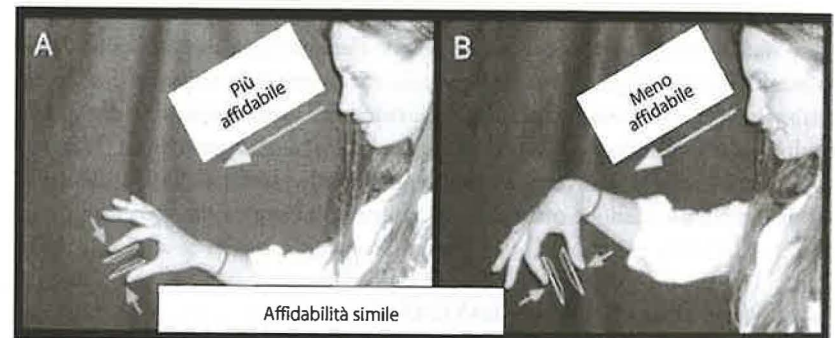



fig. 4.8. Logica dell'esperimento di Gephstein e Banks [2003]. A causa dei fattori proiettivi che vincolano lo stimolo visivo, la stima di una grandezza è più affidabile quando l'estensione da stimare è ortogonale rispetto a quando essa è parallela alla direzione di mira (frecce in alto). La stima tattile dipende invece solo dal grado di apertura delle dita (frecce in basso) ed ha quindi affidabilità simile nelle due direzioni.

fonte: Adattata da immagine di Sergei Gepshtein, per gentile concessione.

diventa invece sempre più schiacciata, fino al caso limite in cui l'estensione è parallela alla linea di mira. In quest'ultimo caso le proiezioni delle due superfici sono di fatto sovrapposte sulla retina e l'informazione visiva sulla loro distanza risulta molto poco affidabile. Nel modello dell'integrazione ottimale, dunque, il peso attribuito al canale visivo dovrebbe essere minimo in quest'ultimo caso e massimo nel primo. Il peso attribuito al canale tattile dovrebbe invece restare invariato o variare molto poco, perché la stima tattile dell'estensione dipende solo dall'apertura delle dita, rilevabile in modo simile in entrambi i casi.

Gephstein e Banks hanno controllato queste previsioni in un esperimento simile al precedente, in cui i volontari dovevano stimare la grandezza di un oggetto solo con la visione, col tatto, o con entrambi i canali, e l'oggetto poteva essere ruotato in diverse maniere rispetto alla linea di mira. Anche in questo caso, le previsioni hanno confermato le previsioni del modello dell'integrazione ottimale: quando la visione era più affidabile, la stima multisensoriale era dominata dal canale visivo, ma quando diventava meno affidabile l'influenza tattile diventava molto più forte.Un aspetto interessante di questi ultimi risultati è che il modello potrebbe in questo modo rendere conto anche dei fenomeni attribuiti al principio della modalità più appropriata, di cui abbiamo anche parlato in precedenza. Implementando un principio di integrazione che tiene conto dell'affidabilità dei segnali unisensoriali, la percezione multisensoriale sceglierebbe, di fatto, la modalità più appropriata per la situazione relativa all'oggetto sotto osservazione. Questo avverrebbe tuttavia non con una preferenza rigida per una certa modalità di senso in una certa situazione, ma in modo flessibile, in funzione dell'affidabilità dei canali sensoriali. A conferma di questa idea, risultati simili sono stati ottenuti anche per quanto riguarda l'integrazione fra vista e udito nella localizzazione di un oggetto nello spazio extrapersonale [Alais e Burr 2004] e fra vista e propriocezione nella localizzazione della propria mano nello spazio peripersonale [van Beers, Wolpert e Haggard 2002]. Nella parte finale di questo capitolo ci occupiamo invece dei problemi che sono stati sollevati in merito al modello dell'integrazione ottimale e dell'approccio bayesiano in generale.

6. I LIMITI DEL MODELLO BAYESIANO

Negli ultimi anni il modello derivato dall'approccio bayesiano ha ricevuto grandissima attenzione da parte della comunità dei ricercatori, e questo per diversi motivi. Innanzi tutto, come abbiamo visto nel paragrafo precedente, il concetto di integrazione ottimale consente di fare previsioni corrette in molte

situazioni di fusione multisensoriale studiate in questa letteratura. Inoltre, lo stesso schema concettuale ha il potenziale per spiegare le apparenti contraddizioni fra i risultati meno recenti, in particolare in riferimento a quelli che sono stati in passato interpretati come dimostrazioni di dominanza di un senso o del principio della modalità più appropriata. Infine, le ipotesi sui meccanismi per modificare i pesi associati ai singoli canali sensoriali sono abbastanza plausibili dal punto di vista biologico. Ad esempio, è stato proposto che l'affidabilità dei segnali unisensoriali potrebbe venire misurata da meccanismi neurali in grado di stimare la variabilità nelle risposte di gruppi di neuroni [Pouget, Dayan e Zemel 2003]. Questa informazione potrebbe essere quindi disponibile in maniera abbastanza rapida per la modifica dei pesi unisensoriali. Inoltre, l'idea che l'esperienza di specifiche associazioni fra informazioni multisensoriali potrebbe modificare le probabilità a priori associate a determinati stati del mondo sembra coerente con quanto succede nei fenomeni di adattamento prismatico (vedi il cap. 3), e in generale con l'evidente capacità di apprendimento esibito dai sistemi percettivi in funzione dei mutamenti delle condizioni ambientali.

A dispetto dei suoi innegabili successi, tuttavia, il modello dell'integrazione ottimale non è privo di problemi. Il nocciolo della difficoltà incontrata da questo apparato teorico è presto detto: tipicamente, l'ambiente offre all'organismo una grande quantità di stimolazione, che viene continuamente registrata dai diversi canali sensoriali. Ora, in molti casi integrare fra loro i diversi canali è la strategia appropriata. In particolare, lo è quando le informazioni relative sono *ridondanti*, vale a dire, si riferiscono alle stesse proprietà dello stesso oggetto, misurate nella stessa unità di misura. Ma come fa il sistema percettivo a sapere che i diversi segnali multisensoriali provengono dallo stesso oggetto? Una possibilità è che questa decisione sia basata su altre proprietà globali della stimolazione, come la **coincidenza spaziale** e la **sincronicità**. L'assunzione che stimoli spazialmente e temporalmente coincidenti provengano dallo stesso oggetto è ecologicamente plausibile, oltre che coerente con i risultati di diverse ricerche. Secondo alcuni risultati sperimentali, infatti, l'interazione fra i segnali sensoriali tende a ridursi se questi sono sfasati più di 100-200 millisecondi [Shams, Kamitani e Shimojo 2002; Slutsky e Recanzone 2001]. Inoltre altre osservazioni, basate di nuovo su un compito di discriminazione di grandezza con stimolazione visiva, tattile, o multimodale [Gepshtein *et al.* 2005] hanno dimostrato che, quando i segnali unisensoriali coincidono nello spazio, la precisione dei giudizi multisensoriali aumenta, come previsto dal modello ottimale. Quando gli stessi segnali provengono da locazioni diverse, invece la precisione multisensoriale si riduce in funzione della distanza.

Problema risolto dunque? Non proprio. Infatti, l'ipotesi secondo cui la percezione multisensoriale degli oggetti dipende dalla coincidenza spaziotemporale, più che risolverlo, sposta il problema da un'altra parte. A questo punto occorrerebbe chiedersi: come fa il sistema percettivo a rilevare che i segnali unisensoriali sono sincroni o coincidenti nello spazio? Come vedremo nelle parti terza e quarta di questo libro, anche la percezione della sincronicità e della coincidenza spaziale sono il risultato di processi multisensoriali complessi [King 2005] e sensibili al contesto [Hanson, Heron e Whitaker 2008; Navarra, Soto-Faraco e Spence 2007]. Finché non saremo in grado di capire in modo soddisfacente la natura di *questi* processi multisensoriali, evidentemente non saremo neanche in grado di capire i processi di integrazione multisensoriale nella percezione di oggetti. E non solo. Nella misura in cui gli effetti contestuali che influenzano la percezione di spazio e tempo potrebbero dipendere da proprietà strutturali dello stimolo legate a loro volta alla presenza di oggetti nel mondo, l'ipotesi rischia di diventare una spiegazione circolare. Inoltre occorre tenere presente un altro problema. Può accadere benissimo che i segnali unisensoriali siano sincroni e provenienti da posizioni molto simili, ma non siano ridondanti – vale a dire, che non si riferiscano esattamente alla stessa proprietà dell'oggetto da percepire. In un utile lavoro teorico su queste problematiche, Ernst e Bülthoff [2004] hanno proposto di riservare il termine **combinazione multisensoriale** per i processi che entrano in gioco in queste ultime situazioni, processi che sarebbero da tenere distinti da quelli di *integrazione* multisensoriale in senso stretto.

L'integrazione, come abbiamo detto, presuppone informazioni unisensoriali ridondanti e riguardanti proprietà che hanno la stessa unità di misura. Nella combinazione multisensoriale, invece, i segnali multisensoriali possono essere *complementari* o anche *in conflitto*. Immaginate di osservare un oggetto come una tazza da caffè mentre la tenete in mano. La mano tipicamente non andrà a collocarsi sulla superficie della tazza direttamente verso di voi (quella in piena vista), ma piuttosto sul lato o sulla parte posteriore (che non vedete). Di conseguenza, il segnale tattile riguarderà una zona dell'oggetto diversa da quella specificata dal segnale visivo. Molto spesso, addirittura, riguarderà una parte diversa dell'oggetto, caratterizzata da una struttura tridimensionale che non coincide per nulla con quella specificata dalla visione. Un esempio è il caso in cui si guarda la parte davanti della tazza mentre immediatamente sul lato se ne tocca il manico. In alcuni altri casi, infine, il segnale tattile riguarderà addirittura un oggetto diverso. Ad esempio, potrebbe succedere che toccando dietro l'oggetto le dita vengano a contatto con cucchiaino, o una zolletta di zucchero (se state prendendo il caffè). In casi come questo i segnali unisensoriali non vanno

integrati, ma semmai combinati in una maniera più sofisticata per ricostruire la struttura complessiva dell'oggetto e le relazioni spaziali fra le sue parti (la tazza e il suo manico), o tenuti separati e attribuiti a oggetti diversi (la tazza e il cucchiaio). La combinazione multisensoriale, così intesa, non può evidentemente essere costituita da una semplice somma ponderata, perché richiede al sistema di prendere delle «decisioni» su cosa va combinato e cosa no, e come.

Non sono ancora disponibili modelli capaci di fare buone previsioni quantitative nei casi che richiedono combinazione multisensoriale. Alcuni ricercatori hanno cercato di applicare l'approccio bayesiano al problema di «decidere» se due segnali unisensoriali vanno attribuiti a fonti distinte, e vanno quindi tenuti separati ed elaborati in modo indipendente, o se i due segnali hanno una fonte comune e vanno quindi combinati in una singola rappresentazione di oggetto [Kording *et al.* 2007]. In buona sostanza, il modello in questione include una seconda componente a priori, denominata **componente a priori dell'interazione** (*interaction prior*), corrispondente alla probabilità che i due segnali provengano da una fonte comune. Con questa aggiunta, il modello prevede molto bene le risposte di localizzazione di un oggetto in presenza di segnali visivi e uditivi spazialmente coincidenti o discrepanti. Va detto, tuttavia, che il successo di questo approccio nel prevedere i risultati di un compito di localizzazione non è troppo sorprendente, dato che la componente a priori che viene aggiunta non fa altro che formalizzare l'idea secondo cui la probabilità di una fonte comune aumenta al diminuire della disparità spaziale fra i segnali unisensoriali. Se questo approccio potrà venire generalizzato con successo ai casi più complessi (e teoricamente interessanti) di combinazione multisensoriale è una questione che verrà plausibilmente chiarita dalla ricerca futura. Un'altra possibilità è che la decisione di combinare o tenere separati i segnali unisensoriali sia affidata a conoscenze di alto livello, recuperate in modo più o meno deliberato dalla memoria a lungo termine. Si tratta di una strategia che assomiglia molto alla posizione di Leibniz riguardo al problema di Molyneux di cui tratteremo nel capitolo successivo. L'utilizzo di conoscenze di alto livello implica un livello di flessibilità che dovrebbe, in linea di principio, consentire di affrontare problemi di combinazione multisensoriali complessi. In una ricerca recente, ad esempio, Helbig ed Ernst [2007] hanno studiato l'interazione fra tatto e vista in una condizione in cui il segnale visivo non corrispondeva alla visione diretta dell'oggetto ma alla visione del riflesso dell'oggetto (e della mano che lo toccava) in uno specchio. A causa delle caratteristiche ottiche dello specchio, questa condizione implica che l'informazione ottica sulla posizione dell'oggetto non corrispondeva alla sua posizione reale (specificata invece dal tatto) ma a una posizione posta

dietro la superficie dello specchio stesso. Veniva dunque generata una larga discrepanza fra le posizioni specificate dalla visione e dal tatto, il che in condizioni normali tende a impedire che i due canali interagiscano. Nel caso dello specchio studiato da questi due ricercatori, invece, il grado di interazione multisensoriale era identico a quello osservato quando i segnali visivi e tattili erano esattamente coincidenti nello spazio (perché i partecipanti vedevano direttamente l'oggetto e la mano anziché vederne il riflesso nello specchio). Il risultato potrebbe dipendere dal fatto che, nel caso dello specchio, i partecipanti «sapevano» che la mano riflessa era la loro e in qualche maniera tenevano conto dell'effetto dello specchio. Le conoscenze sulla natura degli specchi avrebbero quindi la capacità di modulare l'interazione multisensoriale, consentendo di trattare la posizione visiva della mano come se fosse coincidente con quella tattile. Risultati simili sono stati osservati anche in alcuni esperimenti in cui i partecipanti venivano informati verbalmente che due segnali sensoriali provenivano o non provenivano da una fonte comune [discussi in Helbig ed Ernst 2008].

Un ultimo esito possibile per i processi di combinazione multisensoriale, infine, è che i segnali unisensoriali non vengano combinati. In linea di principio, in tal caso possono accadere due cose. La prima è che nei diversi canali sensoriali rimanga un'**informazione separatamente disponibile**, e che questa separatamente possa divenire oggetto di accesso cosciente [Soto-Faraco e Alsius 2007]. Secondo Hillis e colleghi [2002], questo tipo di situazioni testimoniano di una differenza fondamentale fra la fusione di informazioni sensoriali *all'interno* di una singola modalità di senso (ad esempio, fra informazioni visive sulla forma e sulla distanza di un oggetto) e informazioni sensoriali *fra* diverse modalità di senso. Mentre le prime sarebbero obbligatorie, le seconde sarebbero «opzionali». Come dire che non si può fare a meno di fondere forma e distanza, anche se ne risulta un percetto (unisensoriale) paradossale come nel caso del condominio di figura 4.6, mentre in alcuni casi limite si può anche non fondere informazioni provenienti da fonti diverse. Il problema diventa, allora, capire quali siano tali condizioni limite. Intuitivamente, ci si dovrebbe aspettare che riguardino situazioni di conflitto estremo fra le informazioni unisensoriali, come quelle in cui la struttura oggettuale specificata dai singoli canali è troppo diversa per poter essere riconciliata con un'unica interpretazione. Bisogna ammettere tuttavia che una teoria in grado di fare previsioni precise su quando vengano raggiunti questi limiti non è stata ancora formulata. Pensate alle situazioni di conflitto fra vista e udito che abbiamo esaminato nel paragrafo 2 di questo capitolo. Anche in quei casi veniva generato un conflitto fra il segnale visivo (un solo flash di luce) e quello uditivo (due *beep*). Tuttavia la fusione in molti casi avveniva lo stesso,

producendo l'impressione della comparsa di due oggetti. Il problema dei limiti della fusione multisensoriale rimane dunque un problema aperto, che sarà anch'esso presumibilmente oggetto della ricerca dei prossimi anni.

Il secondo esito possibile, in condizioni di conflitto estremo fra i segnali unisensoriali, è che l'interpretazione basata su uno dei canali disponibili eserciti una sorta di «veto» sugli altri, impedendone l'ulteriore elaborazione e l'accesso alla coscienza. In tal caso si dovrebbe verificare una forma di **prevalenza** da parte di quel canale sugli altri. Si tratterebbe tuttavia di un processo diverso da quelli di dominanza visiva studiata in passato in situazione di fusione multisensoriale (par. 3). Come abbiamo visto, infatti, questi ultimi casi possono essere ricondotti al modello dell'integrazione ottimale, ipotizzando che le situazioni in cui i risultati indicano dominanza di un senso siano quelle in cui il senso dominante ha un livello di affidabilità molto maggiore di quelli dominati, ricevendo un peso che lo rende predominante. In questi casi dunque il predominio di un senso sugli altri dipende dalla variabilità relativa dei segnali, e può essere descritto bene nel contesto teorico dell'approccio bayesiano. Una situazione diversa si verifica invece quando l'elaborazione dei due segnali evolve nel tempo, come accade ad esempio durante un'attività esplorativa prolungata, e nel corso di questo processo una delle interpretazioni blocca l'elaborazione di interpretazioni alternative. In tal caso, l'interazione multisensoriale dovrebbe avere una dinamica più estesa nel tempo ed essere caratterizzata da tempi di elaborazione più lunghi.

Alcuni dati a favore di quest'ultima possibilità sono stati riportati, in un recente lavoro, da Bruno e colleghi [2007]. Utilizzando un **cubo di Necker** tridimensionale (fig. 4.9), questi ricercatori hanno studiato l'effetto dell'informazione tattile, acquisita esplorando la struttura del cubo con una mano, sulla probabilità di percepire le due interpretazioni dell'oggetto. In condizioni di visione monoculare e in assenza di informazione tattile, infatti, un cubo di Necker tridimensionale oscilla spontaneamente fra la sua interpretazione veridica (un cubo, appunto) e un'interpretazione illusoria (dovuta all'inversione dell'ordine in profondità dei contorni, che produce la percezione di un tronco di piramide). Riprendendo una vecchia osservazione di Shopland e Gregory [1964], Bruno e colleghi hanno mostrato che l'esplorazione attiva del cubo può bloccare completamente l'insorgere dell'interpretazione illusoria, che è appunto incompatibile con la struttura tridimensionale registrata dal tatto attivo, ma solo in una «finestra di veto» collocata circa un secondo dopo l'inizio dell'esplorazione attiva e solo approssimativamente per un paio di secondi. Una volta trascorso questo periodo, l'interpretazione illusoria può tornare a verificarsi. Il risultato,

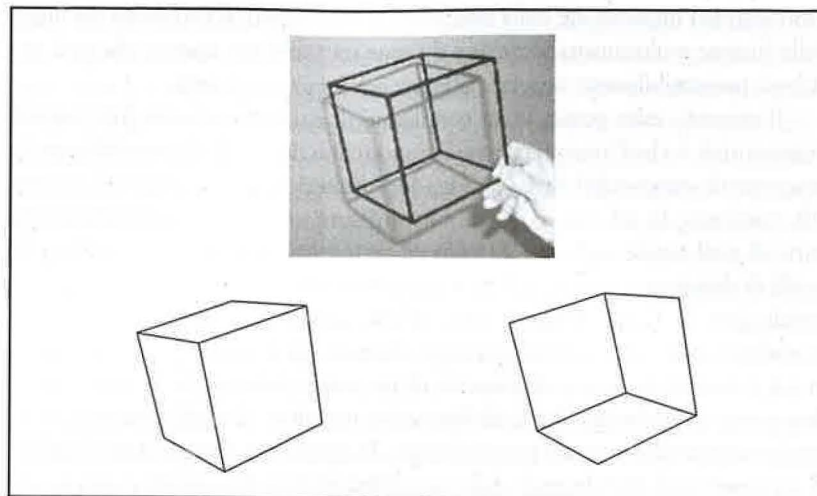



fig. 4.9. Un cubo di Necker tridimensionale. In condizioni di visione monoculare, questo stimolo oscilla spontaneamente fra l'interpretazione veridica (un cubo, sinistra) e l'interpretazione illusoria in cui i contorni invertono l'ordine in profondità (un tronco di piramide, destra).

confermato anche in un lavoro successivo [Bertamini *et al.* 2010], suggerisce che le interazioni multisensoriali possono avvenire anche su scale temporali maggiori di quelle tipiche della fusione multisensoriale immediata, tenendo conto in maniera flessibile dei cambiamenti qualitativi nell'informazione che si rende disponibile durante l'esplorazione prolungata.

CAPITOLO 5

Il riconoscimento crossmodale

Il riconoscimento di oggetti è uno dei temi maggiormente studiati nell'ambito della psicologia cognitiva e delle neuroscienze. Il riconoscimento crossmodale è la capacità di rendersi conto che un oggetto attualmente rilevato da un canale sensoriale è lo stesso oggetto rilevato precedentemente da un canale sensoriale diverso. In questo contesto teorico si collocano tanto l'antico quesito filosofico di Molyneux, quanto il dibattito sui moderni sistemi di sostituzione sensoriale per ciechi.

1. TEORIE DEL RICONOSCIMENTO

La mente umana dimostra una straordinaria flessibilità nel riconoscere oggetti a dispetto di variazioni nella loro grandezza, posizione, orientazione spaziale e nel rappresentare l'insieme di relazioni gerarchiche fra parti che li definisce. Per rendere conto di questa flessibilità sono stati proposti due approcci teorici diversi [Bruno 2004a]: la teoria delle *rappresentazioni strutturali* e la teoria dei *template*.

Secondo la **teoria delle rappresentazioni strutturali** [Biederman 1987], gli oggetti sono rappresentati in memoria da insiemi di componenti primitivi volumetrici, chiamati **«geoni»**. Nel riconoscimento, le rappresentazioni strutturali conservate in memoria vengono confrontate con quelle derivate sulla base dell'informazione sensoriale disponibile. Dato che la rappresentazione strutturale non è metrica ma qualitativa (vale a dire, in termini di relazioni come «sopra-sotto», «vicino-a», e così via), il riconoscimento può avvenire anche in presenza

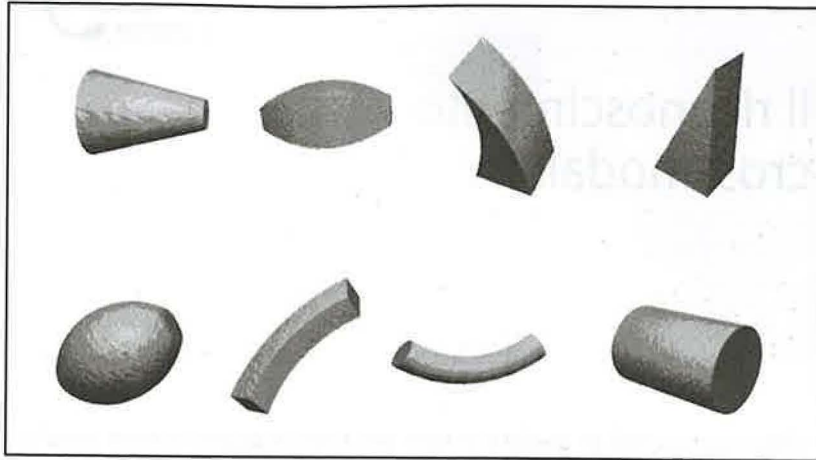



fig. 5.1. Alcuni esempi di primitive volumetriche («geoni») utilizzate per rappresentare la struttura tridimensionale di oggetti. Ad esempio, la rappresentazione di una tazza può essere ottenuta da una combinazione degli ultimi due geoni a destra in basso.

fonte: Michael J. Tarr, Brown University, http://www.tarrlab.org/.

di variazioni nella grandezza delle parti o dell'intero oggetto. Inoltre, dato che viene rappresentata la struttura tridimensionale dell'oggetto, questa rappresentazione rimane uguale anche quando l'oggetto cambia posizione o orientazione nello spazio. Il concetto di geone, derivato da quello di «modello 3D» nei lavori di Marr e Nishihara [1978] e Marr [1982], costituisce il fondamento teorico su cui si regge la teoria. Secondo Biederman, è possibile definire un insieme di soli 24 geoni, ognuno dei quali corrisponde a un volume semplice specificato da proprietà bidimensionali dell'immagine retinica (vedi fig. 5.1). In particolare, sempre secondo Biederman, ogni geone corrisponde alla presenza, nell'immagine retinica, di un insieme specifico di caratteristiche geometriche semplici: per i contorni, ad esempio, se questi sono curvi o rettilinei, paralleli o non paralleli, convergenti o divergenti. Queste caratteristiche sono (salvo casi particolari) indipendenti dalla posizione dell'oggetto rispetto al punto di vista. Ad esempio, un margine oggettuale curvo non produce praticamente mai un contorno rettilineo nell'immagine retinica (dovremmo assumere un punto di vista molto particolare, e improbabile). Secondo Biederman, dunque, il riconoscimento può avere luogo indipendentemente dalle variazioni di posizione rispetto al punto di vista dell'oggetto, perché le caratteristiche grazie alle quali vengono codificati i geoni sono, appunto, ampiamente indipendenti da queste variazioni.

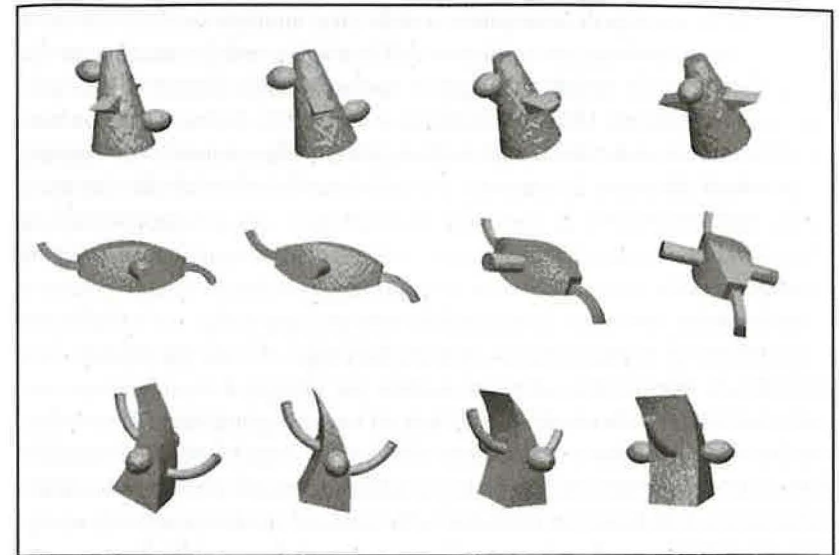



fig. 5.2. Esempi di stimoli sperimentali per lo studio del riconoscimento al variare della posizione rispetto al punto di vista.

fonte: Michael J. Tarr, Brown University, http://www.tarrlab.org/.

La principale difficoltà incontrata dalla teoria delle descrizioni strutturali è di natura empirica. Secondo la teoria di Biederman, infatti, la prestazione umana nel riconoscimento di oggetti dovrebbe essere del tutto indipendente dal punto di vista rispetto al quale l'oggetto viene osservato, a patto che siano disponibili, nell'immagine retinica, le caratteristiche necessarie all'estrazione dei geoni. Considerate un insieme di «oggetti» sperimentali come quelli in figura 5.2, e immaginate un compito di riconoscimento in cui i partecipanti vedono due esemplari in sequenza e devono decidere, il più velocemente possibile, se sono uguali o diversi. Per la teoria di Biederman, il tempo di risposta per emettere correttamente la risposta «uguale» dovrebbe essere indipendente dalla differenza di orientazione fra le due viste dell'oggetto. Ad esempio, nella prima riga, dovrebbe essere ugualmente facile confrontare la prima e la seconda vista o la prima e la quarta vista. La difficoltà è che questa previsione risulta falsificata dai risultati di molti esperimenti: il tempo di risposta varia, e anche molto, in funzione della differenza di orientazione [Tarr e Pinker 1989; Tarr 1995; Tarr e Bülthoff 1995]. Questi risultati hanno portato diversi ricercatori a proporre una teoria basata sul concetto di *template* invece che su quello di rappresentazione strutturale.

Secondo la **teoria dei «template» o delle viste multiple** (*multiple views*), il riconoscimento avviene per confronto dell'immagine retinica attualmente disponibile (la «vista» attuale dell'oggetto) con un archivio interno di viste precedentemente esperite [Bülthoff, Edelman e Tarr 1995]. La memoria non conterrebbe dunque delle descrizioni della struttura tridimensionale dell'oggetto, indipendenti dal punto di vista, ma una collezione di istantanee che catturano come l'oggetto appariva da particolari punti di vista. Queste rappresentazioni funzionerebbero sostanzialmente come delle specie di sagome o mascherine (*template*) a cui la vista attuale viene sovrapposta, per valutare quanto bene vi si adatti (*template matching*). La teoria delle viste multiple spiega bene il fatto che la prestazione di riconoscimento dipenda dal punto di vista, ma incontra una difficoltà che questa volta è di natura teorica: per spiegare il riconoscimento occorre ipotizzare che la memoria contenga un numero grandissimo di *template*, uno per ogni posizione e orientazione dell'oggetto, nonché per ogni possibile alterazione del suo aspetto in seguito a modifiche percettivamente trascurabili. Ad esempio, è un fatto ben noto che l'efficienza nel riconoscimento di un oggetto parzialmente occluso è sostanzialmente equivalente a quello di un oggetto completamente visibile [ad esempio, Bruno, Bertamini e Domini 1997]. Per spiegare questo risultato occorrerebbe immaginare *template* diversi per l'oggetto in piena vista e per tutte le maniere in cui l'oggetto può venire occluso da un altro oggetto.

La contrapposizione fra le due teorie, che ha prodotto un acceso dibattito negli anni '90 del secolo scorso, ha suggerito a diversi ricercatori la possibilità di una teoria ibrida unificata, capace di integrare elementi di entrambi gli approcci. Ad esempio, è stato proposto che il processo di riconoscimento sfrutti simultaneamente più di un meccanismo di elaborazione (**ipotesi delle «multiple routes»**; Vanrie, Willems e Wagemans [2001]). Secondo questa idea, oltre a una via diretta di accesso a un database di viste dell'oggetto, presumibilmente implementato da singoli neuroni nella corteccia inferotemporale [Logothetis, Pauls e Poggio 1995], il sistema potrebbe utilizzare anche una seconda via, indiretta, all'interno della quale verrebbero costruite rappresentazioni strutturali [Corballis 1988]. In alternativa, è stato ipotizzato che il riconoscimento potrebbe sfruttare un meccanismo in grado di ricostruire, grazie a una specie di **interpolazione** fra quelle mantenute in memoria, la vista di un oggetto da un particolare punto di vista – ovviando così alla necessità di averne un numero enorme [Edelman e Poggio 1991; Riesenhuber e Poggio 2000]. Questa seconda ipotesi incorpora anch'essa elementi di entrambi gli approcci, perché l'idea di interpolare fra viste discrete introduce un livello di rappresentazione che non è più solo bidimensionale. Un aspetto interessante, per i fini di questo libro, è che anche nel riconoscimento crossmodale possiamo rintracciare elementi di entrambi gli approcci, come vedremo nell'ultimo paragrafo di questo capitolo. Prima di esaminare la ricerca recente sul riconoscimento crossmodale, tuttavia, ci concederemo una digressione storica. Il problema del riconoscimento di un oggetto da un canale sensoriale all'altro rappresenta, infatti, forse la prima questione di psicologia della percezione multisensoriale che sia mai stata oggetto di un dibattito esteso fra gli studiosi della mente.

2. IL QUESITO DI MOLYNEUX

Uno dei primi pensatori a porre il problema del riconoscimento crossmodale, in epoca moderna, è stato l'irlandese William Molyneux, un filosofo e scienziato dilettante che aveva un interesse particolare nei confronti della percezione e dell'ottica. Nel 1688 capitò fra le mani di Molyneux un riassunto del *Saggio sull'intelletto umano* di John Locke, la cui prima versione sarebbe apparsa a stampa due anni dopo. Stimolato dalla lettura del riassunto, in cui erano trattati problemi che gli stavano a cuore, Molyneux decise di scrivere a Locke per chiedergli se un cieco congenito, avendo imparato a identificare degli oggetti col tatto, sarebbe stato capace di identificarli con la vista nel caso in cui improvvisamente riacquistasse la vista. Nelle parole dello stesso Molyneux: «Supponiamo che un Uomo sia nato cieco, e che avendo nelle mani una Sfera e un Cubo, di circa la stessa grandezza, gli sia stato insegnato quale sia da chiamare Sfera e quale Cubo, finché egli non abbia imparato a distinguerli attraverso la sensazione tattile. Supponiamo ora che entrambe vengano allontanate e poste su un tavolo, e che la Vista gli venga restituita. Chiedo: Sarà egli in grado con la sola vista, e prima di toccarli ancora, di sapere quale sia la Sfera e quale il Cubo? E chiedo inoltre: Sarà egli in grado con la sola vista, prima di allungare in avanti la Mano, di sapere se egli sarà in grado di raggiungerli per afferrarli, ossia di distinguere se siano stati posti a 20 o 1.000 piedi da lui?».

Insomma, Molyneux si chiedeva, e chiedeva a Locke, se l'informazione acquisita col tatto sarebbe stata confrontabile con quella visiva, a dispetto delle differenze qualitative nei due segnali sensoriali e prima che l'ex cieco iniziasse ad avere esperienza delle due assieme. Negli anni immediatamente successivi Locke e Molyneux si scambiarono altre lettere, finché il **«quesito di Molyneux»** finalmente trovò posto nella seconda versione del saggio di Locke, pubblicata nel 1694, balzando all'attenzione della comunità dei filosofi naturali dell'epoca.

La morte di Molyneux, avvenuta pochi anni dopo, gli impedì però di partecipare al dibattito filosofico e scientifico che coinvolse le migliori menti del Settecento, e che continua per molti versi ancora oggi [Dagenaar 1996].

Il quesito di Molyneux pone, sotto forma di esperimento mentale [Jacomuzzi, Kobau e Bruno 2003], una serie di questioni cruciali a proposito dell'interazione fra vista e tatto nella percezione della forma di un oggetto manipolabile. La prima e più ovvia è quella del ruolo dell'esperienza in questi processi multisensoriali. Secondo l'empirista Locke, e anche secondo Molyneux, l'ex cieco non sarebbe stato in grado di identificare il cubo e la sfera, perché questo individuo non ha mai avuto la possibilità di associare le sensazioni tattili collegate ai due oggetti con sensazioni visive. In un approccio strettamente empirista, infatti, il trasferimento dal tatto alla vista diviene possibile solo attraverso questa forma di apprendimento associativo. Si è ritenuto per molti anni che questa posizione empirista trovi conferma nei resoconti di pazienti che hanno recuperato la vista in età adulta dopo la rimozione chirurgica di cataratte congenite [Cheselden 1728; von Senden 1932]. Tuttavia, il decorso postoperatorio che caratterizza questo tipo di pazienti rende in realtà impossibile distinguere fra difficoltà visive dovute a incompleto apprendimento associativo e difficoltà dovute al processo di guarigione dopo l'intervento.

Inoltre, diverse ricerche nell'ambito della psicologia della percezione contemporanea sembrano suggerire che Locke aveva torto. In un classico esperimento Caviness [1964], ad esempio, ha costruito una serie di «oggetti» non familiari, di grandezza e forma simili (avevano tutti la parte posteriore convessa e la parte anteriore con cinque protuberanze variamente disposte). Usando questo materiale sperimentale, Caviness ha poi chiesto a un gruppo di volontari di tastare alla cieca uno di questi oggetti e quindi di scegliere, all'interno della serie completa di dieci che ora potevano vedere ma non toccare, quale oggetto avessero appena tastato. L'accuratezza dei volontari sfiorava il 90%, il che suggerisce che questi erano in grado di confrontare la forma rilevata col tatto con quella rilevata dalla visione, anche se non avevano mai avuto esperienza diretta delle forme utilizzate nell'esperimento. In un'altra celebre ricerca, Meltzoff e Borton [1979] hanno studiato il trasferimento fra tatto e vista a un mese dalla nascita. Questi autori avevano offerto ai bebè dei succhiotti lisci o ruvidi. Dopo questa iniziale fase esplorativa, hanno mostrato ai bebè due sfere, una liscia e una ruvida. I bebè tendevano a guardare più a lungo l'una o l'altra di queste, a seconda dell'esperienza tattile precedente: se avevano tenuto in bocca il succhiotto liscio, guardavano di più la sfera liscia, mentre se avevano tenuto in bocca il succhiotto ruvido, guardavano più a lungo la sfera ruvida. Il dato sembra confermare che la capacità di confrontare la forma rilevata col tatto (nel caso dei bebè, con la bocca) con quella rilevata dalla vista è presente già pochissimo dopo la nascita [ma vedi anche Streri 1987; Streri e Molina 1993]. Il che non significa tuttavia che l'apprendimento non possa avere un ruolo nei processi di percezione multisensoriale: come abbiamo già visto nel capitolo precedente, nel contesto del modello bayesiano dell'integrazione multisensoriale.

Ma la ricchezza teorica del quesito di Molyneux non si esaurisce con il problema dell'empirismo. Un secondo aspetto di grande interesse, ad esempio, è questo: ammesso che vi sia un qualche trasferimento di informazione dalla vista al tatto, qual è la natura del processo coinvolto? In particolare, la condivisione avviene *prima* che i due percetti, tattile e visivo, siano esperiti consciamente, andando a formare un percetto multisensoriale unico e non più analizzabile nelle sue componenti originarie? O piuttosto i due canali sensoriali vengono elaborati fino a produrre due percetti coscienti separati, l'esperienza tattile e l'esperienza visiva, e solo a quel punto interagiscono a un livello postpercettivo, attraverso processi che assomigliano di più al ragionamento e al pensiero logico? Un'idea simile alla prima fu proposta dal filosofo francese La Mettrie [1745], che riteneva possibile che certe proprietà oggettuali (i **sensibili comuni** di Aristotele) venissero rappresentate in un'unica maniera da processi innati, consentendo quindi un immediato trasferimento dal tatto alla vista anche in un ex cieco improvvisamente risanato. Una posizione molto simile alla seconda fu difesa invece dal filosofo tedesco Leibniz [1765], quando suggerì, di nuovo in contrapposizione all'empirismo di Locke, che l'ex cieco di Molyneux avrebbe potuto usare la propria conoscenza di concetti astratti come «superficie curva», «spigolo», e così via, per guidare l'analisi della struttura tridimensionale rilevata con la vista. Come vedremo sempre nell'ultima parte di questo capitolo, ci sono motivi di ritenere che entrambe queste forme di interazione visuotattile possano avvenire, anche se in condizioni diverse.

Un terzo aspetto, infine, riguarda il ruolo dei processi di elaborazione e controllo delle azioni nell'interazione multisensoriale. In un recente lavoro, Ghazanfar e Turesson [2008] hanno suggerito che la questione di Molyneux potrebbe essere in realtà uno pseudoproblema, in quanto la fusione dei segnali sensoriali comporta la costruzione di rappresentazioni multisensoriali di natura essenzialmente motoria, ossia riferite a una specifica azione. Sarebbe dunque solo nel contesto di queste rappresentazioni motorie che la relazione fra informazioni tattili e visive diviene sensata, per cui, «dato che la questione di Molyneux non chiama in causa alcun movimento, va forse considerata una questione mal posta, almeno dal punto di vista neuroscientifico». In realtà, come si legge

nella formulazione originaria della domanda, in testa al capitolo, un riferimento alla potenzialità per l'azione (e alla codifica dello spazio a questo fine) era presente alla fine del quesito. Non c'è dubbio tuttavia che il dibattito filosofico e, fino a pochi anni fa, scientifico, ha dedicato assai poca attenzione al potenziale ruolo dei processi motori nella percezione in generale e in quella multisensoriale in particolare. Non a caso, la seconda parte del quesito di Molyneux non trovò spazio nella trattazione di Locke e rimase del tutto al di fuori del dibattito successivo. Come vedremo fra poco, tuttavia, negli ultimi decenni anche in questo le cose sono decisamente cambiate. La distinzione fra tatto passivo e tatto attivo ha, infatti, assunto un ruolo fondamentale nello studio contemporaneo del riconoscimento crossmodale.

3. LA SOSTITUZIONE SENSORIALE

Le diverse sfaccettature del quesito di Molyneux di cui abbiamo appena discusso non rappresentano unicamente un insieme affascinante di questioni teoriche, ma hanno anche importanti implicazioni pratiche. Prima fra tutte quella di andare al nocciolo del problema della cosiddetta **sostituzione sensoriale**, ovvero della possibilità che individui deprivati di una modalità sensoriale possano utilizzare le sensazioni provenienti dagli altri sistemi sensoriali in maniera vicariante, per sopperire alle informazioni non più disponibili.

Uno dei personaggi più fortunati nel ricco campionario dei supereroi che popolano l'universo dei fumetti può servire bene a introdurre il concetto della sostituzione sensoriale. A causa di un incidente stradale, Matt Murdock diventa cieco ma ben presto scopre che il trauma, nel togliergli la vista, gli ha anche reso straordinariamente sensibili udito, tatto, gusto e olfatto. Grazie a questi supersensi sostitutivi della vista Matt Murdock, divenuto *Daredevil* (nella versione italiana, semplicemente *Devil*) sconfigge criminali e difende gli oppressi compiendo le stesse incredibili imprese di un supereroe vedente che si rispetti. Ad esempio, percepisce a grande distanza il battito cardiaco dei propri nemici, per localizzarli e conoscerne lo stato emotivo, tocca il minuscolo rilievo dei caratteri a stampa su un quotidiano per leggerne le parole, segue la traccia olfattiva del sigaro di un suo nemico per trovarne il nascondiglio nel quartiere newyorkese di *Hell's Kitchen*. Possiamo trascurare qui la questione della plausibilità biologica dei supersensi di Daredevil, che violano molti vincoli fisici e biologici al funzionamento dei recettori sensoriali. Il personaggio ci serve invece per introdurre un'idea che continua a esercitare una grande attrattiva su scienziati e pubblico,

l'idea secondo la quale quando un «senso» viene a mancare è possibile sfruttarne un altro per ottenere le stesse informazioni. L'idea, insomma, che i sensi possano essere sostituiti da altri sensi.

Di per sé, l'idea che alcune informazioni di norma rilevate per mezzo di un canale sensoriale possano essere veicolate anche da altri canali non è particolarmente rivoluzionaria. Come già accennato, Aristotele notava che proprietà come forma, numerosità e movimento sono disponibili alla vista, al tatto e all'udito, il che lo portava a ipotizzare un «senso comune» specifico proprio per queste proprietà (ne abbiamo già accennato nel capitolo 1 e all'inizio di questo). E tutti abbiamo in mente il bastone bianco che i ciechi utilizzano per ottenere, col tatto e con l'udito, alcune informazioni che noi rileveremmo con la vista. O ancora, pensate all'abilità di certe persone sorde nel leggere le labbra. Questa idea, tuttavia, ha preso una forma diversa e assai più ambiziosa a partire dalla seconda metà del secolo passato, quando si è cominciato a esplorare la possibilità di costruire veri e propri sistemi protesici (soprattutto per ciechi, ma non solo) basati sull'idea di inviare a un canale sensoriale diverso l'informazione non più rilevabile da quello mancante. Questi sistemi sono stati chiamati **dispositivi di sostituzione sensoriale**.

I primi sistemi di sostituzione sensoriale erano basati sull'idea di sostituire la visione col tatto, e prevedevano l'utilizzo di apparecchi applicati sulla schiena, o in seguito sulle dita o sulla lingua, per somministrare stimolazioni vibratorie o elettriche codificate da un computer a partire da immagini rilevate da telecamere [Bach-y-Rita 1967]. Questi sistemi hanno dimostrato una qualche utilità nei compiti di discriminazione di forma, in cui l'analogia fra forma visiva bidimensionale e forma tattile è abbastanza naturale [per rassegne recenti vedi Bach-y-Rita 2004; Bach-y-Rita e Kercel 2003], e potrebbero essere utili in alcuni compiti di orientamento spaziale [Segond, Weiss e Sampaio 2005]. Successivamente sono stati proposti sistemi in cui le informazioni rilevate da una telecamera vengono convertite in suoni (sonificazione) secondo regole che l'utente deve apprendere [Arno *et al.* 1999; Cronly-Dillon, Persaud e Gregory 1999; Fontana *et al.* 2002; Meijer 1992; Veraart *et al.* 1998]. Questi sistemi sembrano anche di qualche utilità per compiti spaziali e in particolare per la localizzazione a distanza [Proulx *et al.* 2008]. Sulla base di risultati come questi, i sistemi di sostituzione sensoriale sono stati talvolta pubblicizzati dai media come potenziali soluzioni per il problema della cecità (vedi ad esempio «BBC News» del 7 ottobre 2003).

Tuttavia, se si esamina lo stato dell'arte armati di qualche nozione di percezione multisensoriale, non è difficile rendersi conto che l'entusiasmo nei confronti di questi sistemi protesici è probabilmente eccessivo. Innanzi tutto, dopo

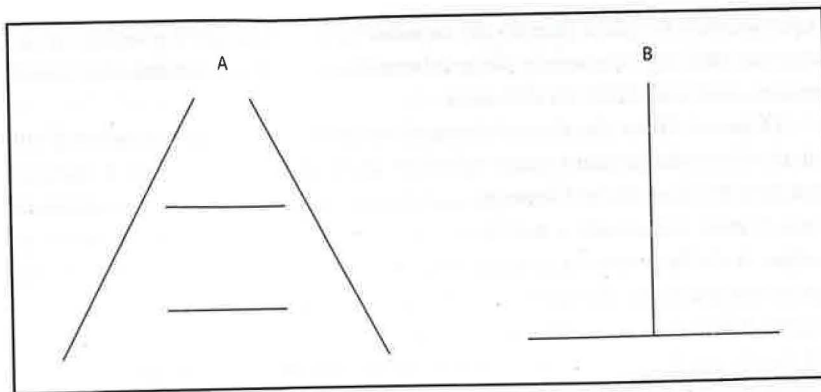



fig. 5.3. L'illusione di Ponzo (A) e l'illusione verticale-orizzontale (B). In A i segmenti orizzontali sono della stessa lunghezza. In B il segmento verticale ha la stessa lunghezza di quello orizzontale.

quasi cinquant'anni di ricerche, non sono ancora disponibili sistemi in grado di fornire un ausilio realmente vantaggioso per un individuo non vedente e, infatti, ben pochi ciechi li utilizzano e per lo più in contesti sperimentali. Posto che l'informatica e la miniaturizzazione elettronica hanno fatto grandi progressi in questo periodo, sembra difficile attribuire questo fenomeno solo a problemi tecnici. Insomma, viene il sospetto che vi sia qualcosa di sbagliato a livello teorico. Inoltre, benché le protesi a sostituzione sensoriale sembrino consentire, dopo sufficiente addestramento, di svolgere almeno alcuni compiti spaziali, questo sembra vero quando a usarle sono persone vedenti (di solito, partecipanti a esperimenti) o individui che sono diventati ciechi in età adulta, ma molto meno vero se l'utente è un cieco congenito. Ad esempio, alcune ricerche sulla percezione di illusioni visive per mezzo di un **sistema di sonificazione** [Renier *et al.* 2005; Renier, Bruyer e De Volder 2006] suggeriscono che nella sonificazione l'illusione di Ponzo e l'illusione verticale-orizzontale (fig. 5.3) non vengono percepite dai ciechi congeniti, ma solo dai controlli vedenti. Può darsi che questo tipo di limitazioni possano venire superate facendo iniziare l'addestramento con questi sistemi poco dopo la nascita [Aitken e Bower 1983], ma stabilirlo con certezza è difficile, anche a causa dei delicati problemi etici che sorgerebbero nella sperimentazione.

In una rassegna di alcuni anni fa, Lenay e colleghi [2003] hanno suggerito che il concetto stesso di protesi basata sulla sostituzione sensoriale sarebbe viziato da un doppio fraintendimento. Secondo questi autori, queste protesi non funzionano grazie a una *sostituzione* né agiscono a livello *sensoriale*. Vale la pena

esaminare questo argomento in dettaglio, anche per le sue implicazioni per la prospettiva multisensoriale che proponiamo in questo libro. Secondo questi ricercatori, in questo tipo di protesi non si può parlare di una vera e propria sostituzione perché non viene riprodotta l'esperienza qualitativa del senso sostituito. E questo non solo perché alcuni aspetti della fenomenologia della visione non vengono codificati (ad esempio, nessun sistema di sostituzione sensoriale può restituire l'esperienza di vedere i colori), ma soprattutto per un motivo più profondo. Infatti, anche le informazioni spaziali convogliate per via tattile o uditiva rimangono del tutto diverse da quelle visive. Più che di sostituzione, si dovrebbe dunque parlare di *aggiunta* di funzionalità a un canale sensoriale intatto: un po' come dire che la protesi non è molto di più che un bastone bianco con un raggio d'azione più ampio e articolato. Inoltre, secondo Lenay e colleghi non si dovrebbe parlare di sostituzione sensoriale ma semmai di sostituzione *sensomotoria*. Infatti, non è affatto chiaro fino a che punto la conoscenza dell'ambiente esterno fornita dalla protesi sia veramente dovuta a processi di codifica sensoriale. Diversi dati inducono invece a pensare che un ruolo cruciale sia giocato dalla covariazione fra azioni esplorative e loro conseguenze sensoriali.

Particolarmente istruttiva, a questo proposito, è la lettura del resoconto pubblicato da Guarniero [1974], un cieco acquisito che si è sottoposto a un addestramento intensivo con il sistema visuotattile di Bach-y-Rita. Guarniero riferisce che, con l'addestramento, forme e posizioni non venivano più percepite sulla schiena (su cui veniva posto l'apparato che somministrava le vibrazioni) ma si collocavano «in uno spazio bidimensionale», che tuttavia «non sembrava essere esterno». A seguito dell'esplorazione dell'ambiente con oggetti posti in diverse posizioni, Guarniero dice di avere imparato in parte a riconoscerli e a valutare le distanze a cui erano posti. Tuttavia, per far ciò era necessario fare dei ragionamenti consapevoli usando indicatori come la grandezza tattile di oggetti familiari o la loro posizione sul piano pittorico. Guarniero dice esplicitamente che «l'utilizzo di tali indici richiedeva una sorta di deduzione cosciente», anche se aggiunge che dopo molti giorni di addestramento non era più necessario «fare attenzione al fatto che stavo facendo tali deduzioni». Perlomeno sulla base di questo resoconto introspettivo, dunque, non è per nulla chiaro se il vissuto fosse simile a quello della percezione vera e propria. Una cosa tuttavia sembra chiara: solo esplorando l'ambiente durante l'addestramento prolungato la capacità di rilevare informazioni spaziali attraverso il sistema diventa ragionevolmente automatica e priva di sforzo.

Tutto ciò non significa però che questo tipo di ricerca non sia interessante: lo è, e lo è moltissimo in una prospettiva multisensoriale. Ad esempio,

lo studio della sostituzione sensoriale è in grado di fornirci informazioni utili sul coinvolgimento di diverse aree corticali nell'elaborazione multisensoriale e sulla loro plasticità [Amedi *et al.* 2007; De Volder *et al.* 1999; Finney, Fine e Dobkins 2001; Poirier *et al.* 2007]. Un altro elemento teorico interessante che emerge dalla ricerca sulla sostituzione sensoriale riguarda proprio il ruolo dell'azione e dell'attività esplorativa nell'apprendimento all'uso del sistema. L'importanza di considerare la percezione come un processo attivo e non passivo è uno degli aspetti centrali della prospettiva multisensoriale proposta in questo libro. La ricerca sulla sostituzione sensoriale fornisce senza dubbio ulteriori elementi a sostegno di questa conclusione: questi sistemi acquistano una qualche efficacia solo nella misura in cui consentono un'efficace esplorazione dell'ambiente, esponendo l'utente alla covariazione sistematica di azioni intenzionali e loro conseguenze multisensoriali [O'Regan e Noe 2001]. Ciò suggerisce anche che uno specifico limite di questi sistemi potrebbe risiedere proprio nell'impossibilità fisica di sostituire pienamente le covariazioni possibili rispetto a un canale con quelle presenti in un canale diverso [Auvray *et al.* 2005].

Il problema delle regole da applicare per ottenere una sostituzione efficace fra diversi canali sensoriali si ricollega a un altro tema che abbiamo affrontato in diversi punti del libro, l'ipotesi di **equivalenza fra segnali sensoriali diversi**. Si tratta di un'ipotesi su cui ci sono state molte discussioni [Marks 1983], in particolare in riferimento ad analogie che sembrano naturali, ma per cui non esistono realmente basi teoriche o empiriche [Kubovy 1988]. Considerate uno dei sistemi di sostituzione visuoacustici discussi sopra [Meijer 1992]. In questo sistema, il suono codifica le proprietà di una matrice bidimensionale di intensità luminose (l'immagine acquisita dalla telecamera) usando la frequenza per la posizione verticale (basse frequenze per la parte bassa dell'immagine, alte frequenze per quella alta) e il tempo per la posizione verticale (per ogni immagine viene prodotto un suono complesso, la cui parte iniziale codifica la parte sinistra dell'immagine mentre quella finale codifica la parte destra). La regola di conversione si adegua all'analogia spazio-frequenza per la dimensione alto-basso, e alla convenzione sulla direzione di lettura per la dimensione sinistra-destra – due principi che sembrano ragionevolmente naturali. Tuttavia l'applicazione di questa regola di conversione è possibile, evidentemente, solo per una singola immagine statica. La possibilità di esplorazione si riduce dunque alla considerazione successiva di una sequenza di «istantanee» sonore, e la struttura temporale del segnale visivo va necessariamente persa. Un problema analogo può essere riscontrato nel caso della sostituzione visuotattile [Bach-y-Rita 1983] dove le variazioni nella stimolazione vibratoria o elettrica incontrano limiti legati alla sensibilità temporale dei meccanorecettori coinvolti.

Insomma, sembra legittimo sospettare che una protesi per ciechi basata sulla sostituzione sensoriale non troverà mai larga applicazione, non perché l'elettronica del sistema non sia sufficientemente veloce o miniaturizzata, ma per il ben più fondamentale motivo che la struttura spazio-temporale dell'informazione ottica conseguente all'azione esplorativa può essere convertita solo in minima parte in struttura spazio-temporale tattile o acustica. Insomma, il problema non è la tecnologia, ma la teoria scientifica che ci sta dietro. Non è un caso, forse, che allo stato attuale gli unici sistemi che hanno trovato utile applicazione in campo biomedico riguardino la **sostituzione vestibolare-tattile** [Danilov *et al.* 2006]. L'informazione rilevata dal vestibolo (ne abbiamo parlato nel cap. 2) riguarda le variazioni di posizione della testa rispetto alla gravità, un segnale codificabile in modo relativamente semplice con variazioni di pressione o di corrente elettrica su una matrice di stimolatori. L'utilizzo di questi sistemi sembra iniziare a trovare una certa diffusione nella riabilitazione di pazienti con deficit vestibolari [Vuillerme *et al.* 2007].

4. RICONOSCIMENTO E ATTIVITÀ ESPLORATIVA

Come abbiamo visto nel capitolo 1, uno dei contributi importanti di James J. Gibson alla teoria della percezione è stato aver chiarito la natura attiva della percezione. I sensi non sono strumenti passivi di registrazione dell'energia nello stimolo, sono sistemi costruiti per esplorare l'ambiente raccogliendo continuamente informazioni. Gibson ha fornito una delle dimostrazioni più convincenti della natura attiva della percezione proprio studiando la percezione tattile degli oggetti. In un famoso esperimento, Gibson utilizzò come stimoli dei semplici stampini da pasticciere, utilizzati per fare i biscotti (*cookie cutters*), e ne studiò il riconoscimento tattile in un compito di identificazione. Gibson confrontò tre condizioni: **tatto passivo** (lo sperimentatore premeva lo stampo contro il palmo della mano del partecipante), **tatto attivo** (il partecipante poteva esplorare attivamente lo stampo) e **tatto dinamico** (lo sperimentatore muoveva lo stampo sulla mano del partecipante, che rimaneva ferma). Il compito consisteva nell'identificare la forma, e poiché questa veniva appresa dai partecipanti vedendo gli stampini, il compito può essere considerato una forma di riconoscimento crossmodale.

Le tre condizioni produssero prestazioni molto diverse: la percentuale di identificazioni corrette era attorno al 50% nella condizione passiva, mentre sfio-

rava il 100% nella condizione attiva. La condizione di tatto dinamico si collocava a metà strada, attorno a circa il 70%. Da questi risultati Gibson concluse che la percezione tattile della forma dell'oggetto sfrutta proprietà strutturali invarianti che si rivelano nel variare temporale dell'informazione (confronto fra tatto passivo e dinamico) e che vengono raccolte da un'attività esplorativa deliberata (confronto fra tatto attivo e tatto dinamico). La conclusione non ci stupisce, perché abbiamo già avuto modo di sottolineare l'importanza dell'azione intenzionale parlando di un oggetto particolare, il nostro corpo. Che il principio valga per gli oggetti in generale sembra del tutto naturale.

Sembra anche naturale attribuire il risultato di Gibson alla percezione e al ricordo di una descrizione strutturale dell'oggetto esplorato, descrizione che può diventare più ricca e articolata grazie all'esplorazione attiva, e che può essere condivisa fra modalità di senso diverse. Gibson riteneva che l'informazione sulla struttura fosse appunto disponibile grazie alle proprietà invarianti sottostanti alla variazione della stimolazione tattile durante l'esplorazione, e sosteneva che invarianti equivalenti possono venire rilevate durante l'esplorazione in un canale sensoriale diverso. L'esperimento di Caviness, che abbiamo descritto parlando del quesito di Molyneux, indicava ad esempio che il trasferimento dalla percezione con tatto attivo al riconoscimento visivo superava il 90% anche con forme non familiari, un risultato coerente con l'ipotesi di Gibson. A ben guardare, tuttavia, questi risultati non sono definitivi perché esiste una possibilità alternativa: l'esplorazione potrebbe consentire la costruzione di una migliore descrizione dell'oggetto, a livello linguistico o semantico. La situazione ripropone dunque il problema del formato della rappresentazione per la percezione di oggetti, problema che abbiamo già visto essere alla base di proposte alternative di soluzione al quesito di Molyneux.

Per studiare questo problema, Reales e Ballesteros [1999] hanno usato una versione multisensoriale di una tecnica che viene chiamata **priming**. Nel *priming*, uno stimolo irrilevante presentato prima dello stimolo sperimentale ne facilita l'elaborazione (in inglese, *to prime* vuol dire preparare, innescare, preavvisare). Nella versione multisensoriale utilizzata da Reales e Ballesteros, i partecipanti dovevano esplorare con il tatto o con la vista un oggetto per 10 secondi (fase di studio), quindi svolgevano un compito distraente per 5 minuti, e infine veniva loro presentato un insieme di 15 oggetti e per ognuno veniva loro richiesto di dirne il nome il più velocemente e accuratamente possibile (fase di test). Per metà dei soggetti le due fasi erano intramodali: ad esempio, se lo studio era visivo anche il test lo era. Per l'altra metà erano invece crossmodali. Inoltre, per metà dei soggetti la fase di studio era tattile mentre la fase di test era visiva; per l'altra metà questo ordine veniva invertito. Infine, e questa è la variabile cruciale, la fase di studio tattile poteva essere di due tipi: nella condizione di codifica semantica, i partecipanti dovevano produrre una frase sensata che includesse il nome dell'oggetto studiato; nella condizione di codifica fisica, dovevano dare un giudizio sul suo volume.

I risultati hanno evidenziato robusti effetti di *priming* sia nella situazione intramodale sia in quella crossmodale, e sia nel caso della codifica semantica sia nel caso della codifica fisica. In altre parole, rispetto alla condizione di controllo senza *prime*, i partecipanti rispondevano più velocemente ma questa facilitazione era simile in tutte le condizioni. Questo dato suggerisce che una codifica semantica del *prime* non è necessaria perché questo abbia un effetto sul riconoscimento dell'oggetto test. E, infatti, Reales e Ballesteros ne hanno tratto la conclusione che la rappresentazione utilizzata per il riconoscimento dell'oggetto è multisensoriale, presemantica, e relativa a una descrizione della sua struttura. Dobbiamo allora concludere che il riconoscimento crossmodale supporta la teoria delle descrizioni strutturali? Non proprio. Nell'esperimento di Reales e Ballesteros non veniva manipolata in maniera esplicita la posizione dell'oggetto dalla fase di studio a quella di test. I dati suggeriscono quindi che la rappresentazione è presemantica e in qualche modo strutturale, ma non ci dicono come venga costruita tale rappresentazione né ci consentono di escludere effetti legati al punto di vista. Una serie di esperimenti condotti da Fiona Newell suggerisce invece che anche nel riconoscimento crossmodale esiste un equivalente dell'effetto del punto di vista.

In un lavoro molto noto [Newell *et al.* 2001], questa ricercatrice ha misurato l'effetto delle variazioni nell'orientazione dell'oggetto sul riconoscimento intramodale visivo o aptico e su quello crossmodale con questi due canali sensoriali. Gli stimoli utilizzati erano forme arbitrarie costruite combinando sei mattoncini del Lego. Quattro di queste forme venivano imparate, in sequenza, mediante la visione o mediante il tatto attivo. Successivamente i partecipanti dovevano svolgere un compito di riconoscimento su oggetti che potevano essere presentati nella stessa orientazione della fase di studio o ruotati di 180 gradi. In questa seconda fase gli oggetti potevano essere presentati nella stessa modalità di senso della fase di studio (riconoscimento intramodale) o nell'altra (riconoscimento crossmodale). I risultati hanno dimostrato che, sia nel riconoscimento visivo sia in quello tattile, la prestazione decadeva quando l'oggetto veniva ruotato rispetto alla fase di studio. Dunque se un oggetto veniva visto in una certa orientazione nella fase di studio, il suo riconoscimento diventava più difficile se nella fase di test questo era ruotato. E la stessa cosa accadeva se la fase di studio

e quella di test era tattile. Tanto nel canale visivo quanto in quello tattile, dunque, emergeva un effetto del «punto di vista». Il risultato più interessate riguarda tuttavia la situazione di riconoscimento crossmodale. In questo caso, infatti, sono emersi dei risultati diametralmente opposti: sia quando la fase di studio era visiva sia quando era tattile, il riconoscimento era migliore quando nella fase di test l'oggetto era ruotato, e peggiore quando era nella stessa orientazione.

Questa sorprendente differenza può essere interpretata facendo riferimento al tipo di azione esplorativa che viene svolta in modo più naturale con gli occhi o con le mani: gli occhi esplorano fissando le parti anteriori dell'oggetto, mentre le mani tendono a esplorare più agevolmente la parte posteriore a causa delle caratteristiche intrinseche dell'articolazione del polso (va precisato, per chiarezza, che in questo esperimento i partecipanti erano lasciati liberi di esplorare gli oggetti come meglio credevano, a patto di non spostarli). Ogni canale sensoriale avrebbe insomma un insieme di «punti di vista» preferenziali, grazie ai quali riesce ad acquisire informazione sulla struttura dell'oggetto in modo ottimale. Quando coinvolge questi punti di vista ottimali, il riconoscimento crossmodale è dunque più efficiente rispetto a quando uno dei due punti di vista è sfavorito. Il risultato è dunque in buon accordo con l'idea che anche il riconoscimento crossmodale utilizzi una codifica legata al punto di vista, e che queste codifiche contribuiscono poi a una rappresentazione della struttura dell'oggetto.

Gli studi della Newell sono anche in buon accordo con le ricerche che hanno esaminato il sistema aptico nella sua funzione esplorativa evidenziando come i movimenti della mano obbediscano a logiche complesse, che dipendono sia dalla forma e dalla capacità di movimento dell'arto sia dallo scopo dell'esplorazione. Lo aveva notato già Gibson, che elencava tre caratteristiche tipiche dei movimenti esplorativi:

1. seguire il margine dell'oggetto con le dita inserendole in eventuali cavità;

2. muovere le dita in modo apparentemente casuale, senza un ordine ovvio;

3. tastare in opposizione fra il pollice e una delle altre dita, talvolta facendole scivolare sulla superficie [Gibson 1966, 125].

Gibson suggeriva che questi movimenti sembrano avere lo scopo di rilevare un insieme di conformazioni tattili (*touch-postures*) discrete, più che una variazione sistematica e costante dello stimolo tattile, e ipotizzava che da queste conformazioni potessero emergere le invarianti strutturali necessarie a rappresentare la forma. Le pionieristiche osservazioni di Gibson sono state confermate e ampliate da una serie di lavori assai interessanti di due altre ricercatrici, Susan Lederman e Roberta Klatzky, che si sono occupate specificamente di studiare le attività esplorative tattili. Nel loro lavoro principale [Lederman e Klatzky 1987], queste ricercatrici hanno descritto una serie di **procedure esplorative**, qualitativamente diverse, ognuna specifica e ottimizzata per un determinato compito. Ad esempio, se sono interessate a valutare il peso dell'oggetto le persone tendono a muovere la mano in verticale (soppesandolo, appunto), mentre se devono identificare l'oggetto tendono a muovere le dita tracciandone il margine. La complessità dei movimenti attuati durante l'esplorazione tattile va dunque valutata con molta attenzione negli studi sul riconoscimento visuotattile, perché potrebbe implicare, in alcune situazioni, differenze fra l'informazione acquisita dai due canali che potrebbero sfuggire al controllo sperimentale [Ernst e Newell 2007].

Va detto, infine, che questi dati a favore di una rappresentazione ibrida della struttura tridimensionale dell'oggetto non escludono la possibilità di un processo di riconoscimento basato su una codifica semantica di più alto livello. Numerosi studi, in passato, hanno studiato il riconoscimento crossmodale visivo di stimoli appresi sotto forma di parole presentate nel canale uditivo [vedi ad esempio Graf, Shimamura e Squire 1985; Roediger e Blaxton 1987]. Questi studi hanno prodotto risultati in parte contraddittori, indicando che in alcuni casi non si osservano effetti di *priming* crossmodale o se ne osservano di ridotti rispetto al caso intramodale. Questi risultati potrebbero tuttavia dipendere dall'utilizzo del materiale verbale. In una ricerca recente Schneider, Engel e Debener [2008] hanno costruito invece un repertorio molto ampio di immagini e suoni riferiti a oggetti naturali. Dopo aver raccolto dati normativi su alcune centinaia di questi stimoli in riferimento al grado di familiarità e alla prestazione tipica nei compiti di identificazione e categorizzazione, questi ricercatori hanno svolto una serie di studi utilizzando un compito di *priming* crossmodale. I risultati hanno confermato l'esistenza di un effetto di *priming* semantico: quando il suono e l'immagine erano semanticamente congruenti le risposte erano sistematicamente più veloci rispetto a quando erano incongruenti.

CAPITOLO 6

La percezione multisensoriale dei cibi

Gli oggetti edibili sono una categoria molto speciale di oggetti, perché fondamentali alla sopravvivenza e profondamente multisensoriali. La scelta se ingerire un cibo e l'insieme delle sensazioni che ne ricaviamo è il risultato di complesse interazioni fra i canali legati a gusto, olfatto, vista, sensi corporei e udito. I principi di integrazione multisensoriale applicabili a questa speciale categoria di oggetti sono tuttavia affini a quelli descritti nei capitoli precedenti. Ciò potrebbe avere implicazioni pratiche importanti per l'industria alimentare.

1. GLI OGGETTI EDIBILI

Come percepiamo gli oggetti dipende anche, in parte, dalla categoria a cui gli oggetti appartengono. Una categoria speciale, per motivi ovvi, è quella degli **oggetti edibili**. Interagendo col mondo veniamo a contatto con oggetti che sono anche potenziali fonti di nutrimento. La capacità di discriminare e riconoscere gli oggetti di questo tipo ha una forte valenza adattiva, e, forse proprio per questo, la percezione degli oggetti edibili è una delle funzioni mentali che più sfrutta tutte le informazioni sensoriali disponibili. La percezione del cibo, infatti, può essere considerata una delle esperienze multisensoriali per eccellenza. La valutazione di una pietanza non si basa unicamente sulle informazioni rilevate dalla bocca e dalla lingua (il «senso» del gusto nell'accezione comune), ma anche da quelle rilevate dall'olfatto. Inoltre viene influenzata dalla vista, dall'udito, e da una serie di altre informazioni di natura uditiva, tattile, somatosensoriale e perfino irritante o dolorifica.

Pensate, ad esempio, al gusto «astringente» che hanno alcuni cibi acidi o i vini invecchiati in botti di legno. L'astringenza è dovuta alla presenza di tannini, sostanze chimiche in grado di produrre effetti tattili sul tessuto interno della bocca ed effetti somatosensoriali a carico dei muscoli della faccia. Per la precisione il gusto dei cibi ricchi di tannini sembra includere una combinazione complessa di tre aspetti: impressione di secchezza delle fauci, impressione di ruvidezza su lingua e palato, e impressione di tensione verso l'esterno dei muscoli delle guance [Lee e Lawless 1991]. Ruvidezza e secchezza, proprietà tattili, vengono percepite muovendo la lingua sulle superfici buccali interessate. La percezione dell'astringenza rappresenta dunque un altro esempio della natura attiva ed esplorativa della percezione, in cui l'attività esplorativa viene svolta dalla lingua, su parti del corpo (il tessuto interno della bocca), ma allo scopo di ottenere informazioni sulla sostanza che si sta per ingerire. O ancora, si pensi all'effetto sul gusto della presenza di anidride carbonica in soluzione, come nelle bevande gassate, e della conseguente sensazione tattile di pizzicore. O infine, pensate all'esperienza gustativa che si può provare con cibi molto piccanti: la sensazione di bruciore del peperoncino, il gusto pungente del pepe, o la sensazione nel naso quando assaggiamo del rafano. Nonostante una grande quantità di studi sulla sensibilità a queste sostanze, e in particolare alla capsicina, la sostanza chimica irritante contenuta nel peperoncino [per una rassegna vedi Green e Lawless 1991], sappiamo ancora molto poco sul ruolo di queste sostanze nella percezione multisensoriale dei cibi. Per questo motivo, in questo capitolo non dedicheremo ulteriore spazio a questi aspetti, per concentrarci invece sulle interazioni multisensoriali fra il gusto vero e proprio e l'olfatto, la vista e l'udito.

Prima di iniziare va fatta però ancora un'ultima premessa. La maggior parte delle persone, almeno in un paese con una diffusa cultura alimentare come il nostro, sa che servire una pietanza alla temperatura sbagliata (troppo calda o troppo fredda), o con la consistenza sbagliata (troppo al dente o troppo cotta) può rovinare il piacere di mangiarla. L'idea che il corretto coinvolgimento di tutti i sensi abbia un ruolo nell'apprezzare il cibo è dunque un'idea abbastanza condivisa, che ha trovato attenzione anche nell'elaborazione culturale più d'élite. Per esempio, l'intellettuale futurista Marinetti scrisse, assieme a un altro artista futurista conosciuto con lo pseudonimo di Fillía, un volume su *La cucina futurista* [Marinetti e Fillía 1932] dove venivano proposti alcuni suggerimenti per aumentare la stimolazione sensoriale durante il mangiare. Questi suggerimenti sottolineano come anche per quegli artisti fosse importante l'interazione tra i sensi. Ad esempio, veniva suggerito: *a*) di curare in modo originale l'armonia della tavola (cristalleria, vasellame, addobbo) coi sapori e colori delle vivande; *b*) l'invenzione di complessi plastici saporiti, la cui armonia originale di forma e colore nutra gli occhi ed ecciti la fantasia prima di tentare le labbra; *c*) l'abolizione della forchetta e del coltello a favore dei complessi plastici che possono dare un piacere tattile prelabiale; *d*) l'uso dell'arte dei profumi per favorire la degustazione: ogni vivanda deve essere preceduta da un profumo che verrà cancellato dalla tavola mediante ventilatori; *e*) l'uso della musica, limitato agli intervalli tra vivanda e vivanda perché non distragga la sensibilità della lingua e del palato e serva ad annientare il sapore goduto ristabilendo una verginità degustativa; *f*) la presentazione rapida tra vivanda e vivanda, sotto le nari e gli occhi dei convitati, di alcune vivande che essi mangeranno e di altre che essi non mangeranno, per favorire la curiosità, la sorpresa e la fantasia.

Le regole della cucina futurista si concentrano sull'aspetto **edonico** (relativo alla possibilità di un'esperienza di godimento) della degustazione, e non poche ricerche (anche per la rilevanza applicativa di questo aspetto del cibo) si sono effettivamente concentrate sul ruolo della multisensorialità nei giudizi di piacevolezza del cibo. I tre paragrafi che seguono, tuttavia, si concentrano più sugli aspetti di base dell'esperienza del cibo piuttosto che sulle componenti edoniche, in cui plausibilmente giocano un ruolo importante anche i fattori culturali. Come vedremo nei paragrafi successivi, molte delle ricerche sugli aspetti di base hanno studiato compiti come ad esempio l'identificazione o il riconoscimento di un cibo, o la rilevazione di un particolare sapore. Questi compiti sono dunque più simili a quelli che abbiamo discusso nei paragrafi precedenti sugli oggetti non edibili. Molte di queste ricerche, inoltre, hanno messo in luce come la percezione del cibo dipenda da processi multisensoriali di base, simili a quelli discussi nei precedenti paragrafi, e non solo da aspetti culturali e dall'apprendimento.

2. INTERAZIONI FRA SEGNALI OLFATTIVI E GUSTATIVI

La maggior parte di noi ritiene che il sapore dei cibi e delle bevande venga percepito all'interno della cavità orale. La storia è più o meno questa. I sapori di base sono quattro, dolce, amaro, acido e salato. Sulla lingua ci sono le papille gustative, distribuite in modo tale da far sì che zone diverse siano specializzate per gustare uno di questi sapori primari: il dolce sulla punta, l'amaro sulla parte posteriore, il salato sui lati nella parte anteriore, l'acido sui lati in quella posteriore. Ogni gusto non è altro che la combinazione di queste quattro categorie di

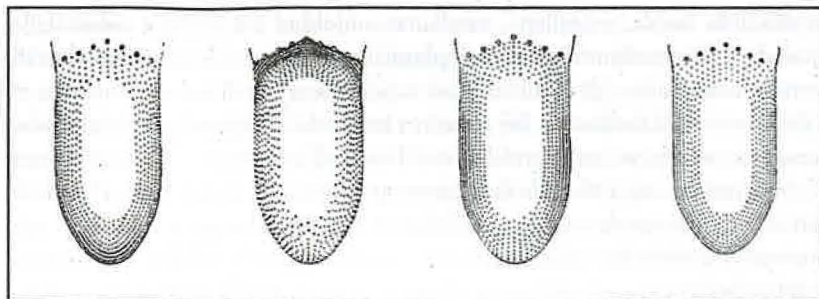



fig. 6.1. La sensibilità (inverso della soglia assoluta) a quattro categorie di sostanza edibile, in funzione della localizzazione sulla lingua, secondo Hanig [1901].

base, i **sapori primari**. Non è forse questo il senso del gusto? In effetti, una storia come questa viene ancora insegnata in molti libri di biologia delle scuole secondarie, e anche in qualche manuale di degustazione nell'industria alimentare. Il problema è che questa storia è falsa. Dipende, infatti, da una serie di distorsioni e fraintendimenti sul significato di osservazioni psicofisiche compiute da un ricercatore tedesco [Hanig 1901] all'inizio del Novecento. Hanig era interessato al rapporto fra superficie linguale e sensibilità, per cui studiò le variazioni della soglia assoluta (minima intensità rilevabile di uno stimolo, in questo caso, chimico) in funzione della posizione sulla lingua, per quattro categorie di sostanza. Le osservazioni di Hanig vengono spesso riassunte utilizzando un diagramma come quello nella figura 6.1, dove la densità dei punti è proporzionale alla sensibilità.

Nel diagramma si può notare che la parte anteriore della lingua risulta leggermente più sensibile al dolce, la parte posteriore all'amaro, e così via. Ma in generale tutta la superficie della lingua risponde a tutte e quattro le categorie di sapore. E in una ricerca più recente, Collings [1974] ha mostrato che le differenze in funzione della localizzazione sono assai piccole, forse addirittura prive di significato pratico. Va rilevato inoltre che il fatto di aver studiato quei quattro tipi di sostanza non implica necessariamente che quelli siano sapori primari. Se Hanig li ritenesse tali non è noto. Ciò che sappiamo è che alcuni ricercatori hanno sostenuto la necessità di affiancare, a questi quattro ipotetici primari, almeno una quinta categoria, il gusto **umami** [Kawamura e Kare 1987]. L'umami, che identifica la sensazione di sapidità che è tipica, ad esempio, di un buon brodo di carne, è considerato da sempre un gusto primario nella cultura giapponese. Sappiamo inoltre che importanti ricercatori contemporanei sui sensi chimici hanno messo in dubbio l'esistenza (nonché l'utilità teorica) dei sapori primari [Halpern 1997], preferendo un modello in cui l'esperienza gustativa risulta

dall'interazione fra meccanismi fisiologici complessi, in cui i quattro (o cinque) gusti primari sarebbero semplicemente punti che definiscono delle polarità in un continuum.

Dunque il gusto è ben più complicato della storia che talvolta si impara nella scuola secondaria. Ma un'ulteriore (e per noi più interessante) complicazione, come abbiamo già detto nel paragrafo 1, è che la percezione dei cibi non dipende solo dal gusto in senso stretto. In realtà, nella percezione del sapore un ruolo fondamentale è svolto dall'olfatto. Provate a fare un esperimento molto semplice: tappatevi il naso e bendatevi gli occhi e poi chiedete a un amico di farvi assaggiare un succo di frutta senza dirvi prima con cosa è fatto. Durante questa prova concentratevi sulle sensazioni che percepite all'interno della vostra bocca. Quelle sensazioni corrispondono al contributo del senso del gusto in senso proprio (inteso come percezione dei famosi gusti primari). Noterete che non siete in grado di riconoscere cosa state bevendo. Ora provate a deglutire il succo che avete tenuto in bocca fino a questo momento e a liberare le narici, espirando. Dopo qualche secondo diventerete coscienti del sapore di frutta che solo allora sarete in grado di riconoscere. Molte persone non riescono a distinguere una mela da una cipolla, o il tè dal caffè durante una prova di riconoscimento con il naso tappato.

Ma non è necessario fare prove come queste per comprendere come spesso siamo vittima di questa **confusione gusto-olfatto**. Molte persone credono di aver perso il senso del gusto quando hanno il naso congestionato a causa di un raffreddore. Ad esempio, non distinguono una mela da una patata. In realtà, sono gli odori che non riescono a raggiungere i recettori nasali all'interno del nostro naso. Dunque quello che noi chiamiamo il gusto di un cibo o di una bevanda è in gran parte dovuto all'olfatto [Murphy, Cain e Bartoshuck 1977]. Per esempio, il gusto del limone è tipicamente descritto come una combinazione di amaro, dolce e acido. In realtà, il caratteristico sapore di limone è dato da particelle di terpeni, biomolecole che raggiungono la mucosa olfattiva attraverso la via retronasale, dalla cavità orale a quella nasale. L'olfatto è dunque un canale sensoriale che ha una doppia funzione, rivolta sia a oggetti odoranti nell'ambiente sia a composti volatili presenti nella bocca [Rozin 1982], e agisce in stretta collaborazione con il gusto in senso proprio costituendo un vero e proprio sistema percettivo specializzato per la percezione del cibo, come aveva ipotizzato già Gibson [1966] (vedi cap. 1).

In questo sistema, tra gusto e olfatto avviene una vera e propria fusione multisensoriale che, come in molti dei casi che abbiamo discusso nel capitolo 4, può addirittura modulare la sensibilità al livello dei processi percettivi di

base. Dalton e colleghi [2000], ad esempio, hanno condotto una serie di esperimenti nei quali alle persone veniva chiesto se riuscivano a percepire l'odore di mandorla (benzaldeide) quando era presentato loro a varie concentrazioni. Usando opportuni metodi psicofisici, Dalton e colleghi hanno trovato la minima concentrazione di benzaldeide necessaria per percepire l'odore (quella che viene chiamata soglia di detezione nasale). Stabilita la soglia, Dalton e colleghi passavano alla fase sperimentale vera e propria. In questa fase, osservarono che le persone erano capaci di sentire l'odore di mandorla a un'intensità dello stimolo molto più bassa della soglia nasale (cioè a una minor concentrazione di quella solitamente necessaria per percepire lo stimolo) quando contemporaneamente essi potevano «gustare» una quantità ugualmente sottosoglia di soluzione zuccherata (contenente saccarina) sulla lingua. Questo risultato indica dunque che la sensibilità del sistema, in condizioni multisensoriali, è migliore di quella in condizioni unisensoriali – un esempio di integrazione «ottimale» nel senso introdotto nel capitolo 4.

Va notato tuttavia che questi effetti non si osservano con tutte le combinazioni di stimoli olfattivi e gustativi. In un esperimento successivo, Dalton e altri hanno riferito che la percezione nasale non era influenzata dal fatto che le persone tenessero sulla loro lingua soluzioni sottosoglia di glutammato monosodico (MSG) o di acqua deionizzata (che non aveva alcun gusto). Questo risultato mostra che gli effetti da aumento multisensoriale funzionano solo per certe combinazioni di stimoli olfattivi e gustativi, in particolare per quelle combinazioni di sapori e odori che normalmente si presentano assieme. Per esempio, come la saccarina e l'aroma di mandorle (benzaldeide) che noi possiamo incontrare in molti dolci e biscotti. Dunque vi sarebbe una stretta relazione tra gli effetti di integrazione multisensoriali e le associazioni che noi apprendiamo.

3. IL RUOLO DEI SEGNALI VISIVI

La nostra percezione del cibo non deriva solo dai segnali sensoriali provenienti dai recettori nel naso e nella bocca. L'esperienza del cibo coinvolge potenzialmente tutti i canali sensoriali, e molti ricercatori si sono sforzati di identificare il contributo di altre informazioni nella percezione del cibo. In particolare, ci si è interrogati su quale sia il ruolo delle informazioni visive sulla percezione del sapore di ciò che mangiamo e beviamo. Un esempio molto convincente è stato riportato da Wheatley [1973]. Questo ricercatore ha descritto la situazione in cui un gruppo di persone era stato invitato a una cena in cui il colore delle pietanze era mascherato da una luce particolare. Le persone iniziarono a mangiare, mostrando di gradire le pietanze. Ma a un certo punto vennero ristabilite le condizioni normali di illuminazione e le persone si accorsero che i cibi che avevano fino a quel momento gradito avevano colori bizzarri: la bistecca era di colore blu, le patate fritte erano verdi, e i piselli erano rossi (fig. 6.2). Wheatley riferisce che la sola visione di quegli «strani» colori fu sufficiente per indurre uno stato di nausea in molte delle persone presenti, che immediatamente si recarono alla toilette sentendosi male. Questo resoconto, seppure aneddotico, suggerisce quanto possa essere potente il ruolo delle informazioni visive nella nostra risposta al cibo.

fig. 6.2. Immaginate di assaggiare questa bistecca con patate fritte e piselli a occhi chiusi. Dopo alcuni bocconi aprite gli occhi e vi rendete conto dei colori. Come reagireste?

fonte: Per gentile concessione Cristy Ho.

L'episodio descritto da Wheatley si riferisce a una reazione di disgusto successiva al momento in cui il cibo era stato assaggiato. Ma è possibile uno specifico **effetto visivo** sul vero e proprio sapore di un cibo? Ad esempio, il sapore di una bevanda cambia se si cambia il suo colore? I risultati di numerosi studi suggeriscono, in effetti, che il colore può avere un effetto potente sulla risposta percettiva a un alimento. Ad esempio, immaginate di chiedere a delle persone se una sostanza inodore profuma. Ovviamente dovrebbero dire di no, giusto? Giusto, ma solo se la sostanza viene presentata in una soluzione trasparente, senza alcun colore. Se la soluzione invece è colorata, molte persone tenderanno a dire che un qualche profumo c'è. Inoltre, se si aggiunge del colore a una soluzione aromatizzata, l'intensità dell'aroma percepito aumenterà. In generale, l'aroma percepito di una sostanza edibile aumenta all'aumentare dell'intensità del colore. Se a una bevanda al sapore di menta si aumenta l'intensità della sua colorazione verde, crescerà anche la valutazione dell'intensità del suo aroma. Ed

è stato dimostrato che l'intensità dell'aroma di fragola di uno yogurt è anch'esso determinato dall'intensità del colore rosso utilizzato per realizzarlo, anche quando il sapore viene valutato da un gruppo di assaggiatori esperti [Teerling 1992].

Il colore può essere usato non solo per cambiare l'intensità percepita del sapore di cibi e bevande, ma anche per modificare la percezione stessa del sapore. In uno studio classico e spesso citato, DuBose, Cardello e Maller [1980] presentavano ai partecipanti al loro esperimento bevande che potevano essere colorate in maniera appropriata (per esempio, bevande alla ciliegia colorata di rosso), non appropriata (per esempio bevanda alla ciliegia colorata di verde) o bevande presentate in soluzioni incolori. I risultati hanno mostrato che l'identificazione del sapore di molte delle bevande era guidata dal colore piuttosto che dal sapore (inteso come effetto congiunto dei segnali gustativi e olfattivi). In particolare, assaggiando una bevanda aromatizzata alla ciliegia la maggioranza dei partecipanti percepiva correttamente il sapore quando questa era di colore rosso, ma molti rispondevano lime quando questa era colorata di verde, e circa il 20% rispondeva arancio quando il colore della bevanda era arancione.

Anche il piacere che noi traiamo da ciò che mangiano può essere influenzato dall'aspetto visivo del cibo. Ad esempio, le persone apprezzano di più gli odori presentati in soluzione, come per esempio l'odore di frutta, se ritengono che il colore della soluzione sia quello appropriato. Christensen [1983] ha dimostrato che le persone ritengono che gli alimenti di consumo quotidiano (come la margarina, il succo di frutta all'arancia, la pancetta o il formaggio) siano di qualità migliore e abbiano un aroma più intenso quando questi alimenti sono colorati normalmente rispetto a quando sono colorati con tinte non appropriate. A proposito della margarina, in particolare, c'è una storia interessante. Quando la margarina venne introdotta per la prima volta in Australia e negli Stati Uniti, le associazioni dei produttori di burro si batterono per evitare che i produttori aggiungessero del colore giallo alla loro margarina. Per questo, le prime margarine prodotte avevano un colore biancastro che le faceva assomigliare al lardo, e questa probabile associazione visiva influenzò molte persone, decretando un insuccesso iniziale del prodotto. La margarina che trovate oggi al supermercato ha un colore più simile a quello del burro, ma è comunque meno gialla. E infatti Louis Chesking del Color Research Institute ha dimostrato che è possibile trarre in inganno le persone semplicemente agendo sul colore dei due prodotti: presentando un burro molto bianco e una margarina molto gialla, il 95% delle persone scambiano il burro per margarina e viceversa [riportato da Packard 1957]. Questo esempio mostra ancora una volta come le informazioni visive possano influenzare drammaticamente la percezione del sapore di quello che mangiamo o beviamo.

Molti degli effetti relativi all'influenza delle informazioni possono essere dovuti alle **aspettative** create dalle informazioni visive sul sapore che cibi e bevande possono avere. Per esempio, molto spesso ci si aspetta che un vino abbia un determinato sapore in funzione del suo colore. Un vino bianco viene di solito indicato come dolce e fruttato, un vino rosé spesso ha un sapore di frutti di bosco, come ribes o lamponi, mentre un vino rosso più invecchiato ha spesso un aroma di frutta appassita. L'accettabilità del bouquet di un vino è significativamente correlata con il suo colore, per cui se si maschera il colore del vino (semplicemente chiedendo alle persone di berlo a occhi chiusi), questa operazione può alterarne l'aroma percepito. Per esempio, le persone difficilmente riescono a individuare il colore del vino dalle informazioni olfattive e gustative che possono ricavare assaggiandolo a occhi chiusi [Sauvageot e Chapon 1983]. Nonostante le informazioni visive abbiano un ruolo così importante anche nella percezione del vino, ci si potrebbe aspettare che esperti assaggiatori di vino, come i sommelier, vengano molto meno influenzati dal colore del vino quando devono giudicare il bouquet del vino stesso. Tuttavia Morrot, Brochet e Dubourdieu [2001], in Francia, sono riusciti a ingannare più di 50 studenti di un corso universitario per esperti di vino. Infatti, i risultati del loro studio mostrarono come i partecipanti venissero indotti a credere che stessero assaggiando del vino rosso, semplicemente colorando di rosso un vino bianco. I partecipanti riferivano che il vino bianco colorato di rosso aveva il bouquet di un vino rosso, e non utilizzavano gli aggettivi che invece avevano utilizzato per descrivere in precedenza lo stesso vino bianco di cui non era stato alterato il colore.

Fino a ora abbiamo dimostrato quanto sia robusta l'influenza del colore sulla percezione del sapore. Ma qual è il meccanismo che permette a ciò che vediamo di alterare ciò che odoriamo o gustiamo? E inoltre, dove avviene nel nostro cervello questa interazione? Molte ricerche hanno mostrato come una delle aree più importanti per l'integrazione di stimoli che derivano dai nostri occhi, dal nostro naso e dalla nostra bocca potrebbe essere la **corteccia orbitofrontale** [Rolls e Baylis 1994]. I neuroni di quest'area rispondono a stimoli olfattivi piacevoli (come l'odore di vaniglia), a stimoli gustativi (come il glucosio), e anche a stimoli visivi che hanno un ruolo critico nel determinare la nostra percezione complessiva del cibo [Rolls 1997]. Robert Osterbauer e colleghi dell'Università di Oxford [2005] hanno utilizzato una tecnica di neuroimmagine, la risonanza magnetica, per indagare come l'attività nella corteccia orbitofrontale cambi in funzione di particolari combinazioni tra odori e colori che essi presentavano

alle persone. Ai partecipanti a questi esperimenti veniva presentata una varietà di odori che includeva quello di limone, di fragola, di menta e di caramello. A volte gli odori venivano presentati da soli, altre volte venivano presentati assieme a particolari colori. A volte il colore corrispondeva all'odore presentato, per esempio, quando il giallo veniva presentato assieme all'odore di limone, mentre altre volte la corrispondenza non c'era – per esempio il colore marrone poteva essere presentato assieme all'odore di fragola. Gli autori hanno osservato che la presentazione di combinazioni appropriate di odori e colori causava un aumento dell'attività cerebrale nella corteccia orbitofrontale dei partecipanti all'esperimento, rispetto a quando gli odori venivano presentati da soli. Al contrario, l'attività cerebrale in questa stessa regione diminuiva quando l'associazione tra colore e odore era incongruente.

Questo studio ci dà dunque indicazioni importanti sui meccanismi neurali sottostanti all'influenza della visione sulla percezione dei sapori. Altri studi di neuroimmagine hanno dimostrato come le persone possano avere la consapevolezza che un odore è familiare, pur senza riuscire a identificarlo. Si tratta di un fenomeno che viene chiamato «sulla punta del naso» (parafrasando il più noto fenomeno linguistico detto «sulla punta della lingua») e che può essere studiato analizzando l'attività cerebrale delle persone nel momento in cui queste stanno cercando di identificare l'odore delle sostanze somministrate [Qureshy *et al.* 2000]. I risultati di queste ricerche dimostrano che una parte del cervello solitamente associata alla sola elaborazione visiva (cioè la **corteccia visiva primaria**) è attiva anche durante questo compito. Questi risultati suggeriscono dunque che probabilmente quando cerchiamo di identificare un odore attiviamo anche dei processi visivi, forse simili a quelli coinvolti nell'immaginazione visiva. Quindi il ruolo della corteccia visiva primaria ci suggerisce un altro elemento per comprendere il perché quello che vediamo ha un effetto così drammatico su quello che odoriamo. Studi di neuroimmagine come quelli appena descritti mostrano quanto sia forte il legame tra vista e olfatto nel cervello, e possono spiegare, a livello neurale, perché quello che vediamo ha un così profondo impatto su quello che odoriamo o gustiamo.

Concludendo questo paragrafo sul ruolo delle informazioni visive nella percezione del sapore del cibo, vogliamo sottolineare una peculiarità dell'interazione tra stimoli visivi e stimoli olfattivi e gustativi. Molte ricerche hanno mostrato come la coincidenza spaziale e la sincronia temporale abbiano un ruolo importante nella fusione multisensoriale. Ma l'influenza visiva sulla percezione del sapere ha luogo molto prima che noi assaggiamo il cibo, e a dispetto di una posizione spaziale differente (il cibo viene visto fuori dalla bocca ma assaporato

all'interno di essa). Il processo multisensoriale coinvolto in questa situazione assomiglia dunque di più alle situazioni che abbiamo discusso parlando del riconoscimento crossmodale, o alle situazioni di fusione in cui i segnali multisensoriali sono disponibili e vengono elaborati nel corso di un'attività esplorativa che si estende nel tempo (vedi cap. 5).

4. IL RUOLO DEI SEGNALI UDITIVI

Mentre le informazioni visive sono potenzialmente presenti prima di portare alla bocca cibi e bevande, quelle uditive tipicamente si rendono disponibili quando addentiamo e mastichiamo. Pensate al rumore quando addentiamo una bella mela croccante, al rumore di un grissino che si rompe fra i denti, o allo scroscio di un liquido portato alla bocca con un bicchiere o un cucchiaio (rumore che, nella nostra cultura, non viene considerato segno di buona educazione, ma che è invece assolutamente corretto in altre). Si tratta, evidentemente, di informazioni di natura ben diversa da quelle fornite dalla visione, ma ricerche recenti sembrano indicare che anche queste possono influenzare la percezione del cibo. Ce ne occupiamo in quest'ultimo paragrafo.

Cibi differenti producono suoni diversi quando li mangiamo. Per esempio, in uno dei primi studi sull'argomento Drake [1963] ha dimostrato come il suono prodotto mordendo e masticando differenti alimenti variasse per ampiezza, frequenza e caratteristiche temporali. Queste differenze nelle caratteristiche del suono possono dare informazioni sulla consistenza dell'oggetto edibile (ad esempio, per stabilire se un frutto è maturo o ancora acerbo), ma anche contribuire all'esperienza di gradimento del cibo stesso. Uno degli aspetti del cibo che più chiaramente sembra contribuire a questi giudizi di gradevolezza è la **croccantezza**. Dall'analisi delle caratteristiche uditive di differenti cibi, infatti, è emerso che i cibi più croccanti producono suoni più intensi quando vengono addentati. Tuttavia, il ruolo delle informazioni uditive nella valutazione degli alimenti sono state investigate sempre adoperando cibi differenti, che avevano sicuramente un aspetto diverso e potevano quindi anche produrre sensazioni tattili non confrontabili all'interno della cavità orale. In questi studi, quindi, è difficile distinguere il contributo della modalità uditiva da quello fornito dalle altre modalità nella percezione del cibo. In una ricerca relativamente recente, invece, Zampini e Spence [2004] hanno mostrato che è possibile influenzare la valutazione di un alimento modificandone solo il suono e mantenendo le altre caratteristiche costanti.

Per il loro studio, Zampini e Spence hanno scelto un prodotto industriale, le patatine *Pringles*. Questo perché queste patatine sono praticamente tutte uguali ed è impossibile distinguerle, sia dall'aspetto sia dal gusto in bocca. In questo modo è quindi possibile manipolare solo il suono prodotto masticando, mantenendo invariate le altre caratteristiche, visive, tattili, olfattive e gustative. I partecipanti erano invitati a dare un unico morso, con gli incisivi, a un gran numero di patatine (per la precisione, 180), tenendo la bocca posizionata vicino a un microfono. Dopo ogni morso dovevano giudicare quanto queste patatine erano croccanti e fresche. Le condizioni dell'esperimento prevedevano nove situazioni diverse. I partecipanti all'esperimento potevano udire il suono reale (senza alcuna alterazione delle frequenze) prodotto dal morso; il suono prodotto poteva essere attenuato o amplificato di 12 dB solo nelle alte frequenze (tra i 2 e i 20 kHz); oppure l'intera gamma del suono poteva essere attenuata di 20 o di 40 dB.

I risultati di questo studio hanno dunque dimostrato che la valutazione dei partecipanti veniva influenzata dal suono. Infatti, le patatine venivano giudicate come più croccanti e più fresche quando il suono complessivo era più intenso oppure, quando solo le alte frequenze (2-20 kHz) erano amplificate selettivamente. Dato che le patatine utilizzate in questo esperimento erano tutte simili tra di loro per quello che riguarda gli aspetti visivi, tattili e legati al sapore, l'unico elemento che variava durante l'esperimento era quello del suono. Inoltre, a ulteriore conferma del ruolo dell'udito nella percezione delle patatine, la grande maggioranza dei partecipanti (ben 15 su 20) riferì che pensava che le patatine fossero state scelte da confezioni diverse, esprimendo sorpresa quando lo sperimentatore spiegò loro che provenivano tutte dalla stessa confezione.

CAPITOLO 7

La sinestesia

In alcuni individui la stimolazione di un canale sensoriale evoca, in modo automatico e immediato, l'esperienza di sensazioni normalmente evocate da altri canali, senza che vi sia effettiva stimolazione di questi ultimi. Questo esempio particolare di multisensorialità è la sinestesia: un fenomeno un tempo ritenuto raro e oggi riconosciuto come più comune, che offre una prospettiva ulteriore per cercare di comprendere i meccanismi alla base delle interazioni stabili fra i sensi.

1. UN FENOMENO CURIOSO

Il termine **«sinestesia»**, dal greco *syn* (unione) e *aisthesis* (percezione), si riferisce a una curiosissima anomalia nella percezione multisensoriale. Negli individui con sinestesia, la stimolazione di un canale sensoriale evoca, in modo automatico e immediato, l'esperienza di sensazioni normalmente evocate da altri canali, senza che vi sia effettiva stimolazione di questi ultimi. Questa esperienza è genuinamente percettiva, non meramente immaginata o pensata. Ad esempio, in un saggio divulgativo di alcuni anni fa uno dei padri della ricerca contemporanea sulla sinestesia, Richard Cytowic, ha discusso il caso di M.W., un individuo che, quando assaporava una pietanza, contemporaneamente sentiva la «forma» del suo gusto in una serie di sensazioni tattili – sensazioni reali, come se degli oggetti gli venissero «sfregati sulla pelle del volto e delle mani» [Cytowic 2002, 4]. M.W. usava la sua sinestesia in modo creativo: Cytowic riporta che per M.W. cucinare voleva dire «creare una pietanza con una forma interessante» e che per esemplificare, M.W. diceva che «lo zucchero rende il gusto "rotondo"

mentre il limone aggiunge delle "punte" al sapore» [*ibidem*, 66]. Anche se alcuni autori la descrivono come una condizione di interesse neurologico o comunque anomala, la sinestesia non è citata nei testi di psichiatria come una patologia mentale e non ha di solito effetti negativi per la vita del sinesteta. La sinestesia va quindi considerata una variante nell'esperienza percettiva di alcuni individui, da accostare probabilmente a fenomeni come l'orecchio assoluto in campo musicale-acustico, o la capacità di discriminazione e identificazione degli odori da parte dei «nasi eccezionali» di certi esperti di profumi.

Descritta per la prima volta – a quanto ne sappiamo – alla fine del Settecento da John Locke nel suo *Saggio sull'intelletto umano*, la sinestesia è tornata gradualmente all'attenzione dei ricercatori nella seconda metà del secolo scorso, inizialmente a proposito dell'associazione colore-tono. Simpson, Quinn e Ausubel [1956] hanno descritto esperienze di luci colorate evocate da stimoli uditivi in età evolutiva. Queste osservazioni sono state confermate in individui adulti da Concepcion Perez [1969], Marks [1975], Jacobs e colleghi [1981] e infine da Cytowic e Wood [1982a] in un lavoro pionieristico in cui veniva descritto il caso di una donna per cui certi suoni evocavano luci colorate e veniva presentato per la prima volta il caso dell'uomo «che gustava le forme» a cui abbiamo accennato sopra. Questi resoconti hanno stimolato alcuni primi sforzi teorici [Cytowic e Wood 1982b; Baron-Cohen *et al.* 1993] e, con il nuovo secolo, un interesse sempre maggiore nei confronti della sinestesia. Da un paio d'anni è in rete una batteria di test per la diagnosi di sinestesia [Eagleman *et al.* 2007] e nel solo periodo dedicato alla scrittura di questo libro sono già uscite diverse decine di nuovi lavori sull'argomento. L'impressione condivisa è che un fenomeno così curioso abbia la potenzialità di darci informazioni preziose sulla natura della percezione multisensoriale e della sua base neurale. Siccome le caratteristiche fenomeniche che sono tipicamente coinvolte nelle **esperienze sinestetiche** sono proprietà che rileviamo nell'interazione con gli oggetti, come colori, forme, toni, sapori, sembra inoltre plausibile che queste potenziali informazioni debbano riguardare in modo particolare i problemi di cui ci siamo occupati in questa parte del libro.

2. UNA PRECISAZIONE TERMINOLOGICA

Innanzi tutto, sgombriamo il campo da possibili equivoci. Il termine sinestesia, soprattutto in passato, è stato spesso utilizzato per fare riferimento al concetto generico di unione fra i sensi, o alle cosiddette **metafore sinestetiche** nel linguaggio. Numerosi esempi di metafora sinestetica sono contenuti ad esempio

negli straordinari *Ossi di Seppia* di Eugenio Montale. Pensate a *I Limoni*, dove un colore ci dà, metaforicamente, una canzone, che a sua volta viene avvertita come una sensazione tattile: «Quando un giorno da un malchiuso portone / ci si mostrano i gialli dei limoni / in petto ci scrosciano / le loro canzoni». Altrettanto numerosi esempi si trovano nella terminologia un po' esoterica dei profumieri o dei sommelier, per cui un vino può essere rotondo, vellutato o spigoloso, o un profumo fresco, dolce, vibrante. La sinestesia di cui parliamo qui non ha nulla a che fare con questo uso della metafora, che ha lo scopo di richiamare alla memoria dell'interlocutore esperienze sensoriali con cui si vuole sollecitare un confronto. I sinesteti di cui parliamo qui hanno invece vere e proprie percezioni, concrete esperienze di caratteristiche fenomeniche come colori o forme in assenza di stimoli fisici corrispondenti, o perlomeno, in quella che potremmo chiamare una forma debole di sinestesia (vedi la fine del prossimo paragrafo), la concreta e unanimemente condivisa esperienza fenomenica della somiglianza o associazione fra due segnali sensoriali diversi.

Lo stesso discorso va fatto inoltre per le cosiddette esperienze sinestetiche che sarebbero associate all'assunzione di sostanze psichedeliche, come l'LSD e la mescalina. Anche se in questo caso alcune persone riportano vere e proprie percezioni indotte da stimoli in altre modalità di senso, la natura di queste esperienze fenomeniche sembra molto diversa da quella dei fenomeni sinestetici in senso stretto. Infatti, le allucinazioni indotte da sostanze psichedeliche sono tipicamente piuttosto complesse e possono includere persone, animali fantastici, giochi di luci e colori. Inoltre le esperienze indotte da droghe sono sempre diverse: talvolta la musica evoca luci, altre volte immagini di animali, e così via. Le esperienze sinestetiche di cui parliamo qui, come vedremo, sono invece semplici e regolari: forme, lampi colorati, caratteristiche della trama, e sempre nelle stesse associazioni. Incidentalmente, questo argomento viene di solito utilizzato per mettere in dubbio la plausibilità di uno dei tentativi di spiegazione della sinestesia, quello secondo cui il fenomeno dipenderebbe dalla disinibizione di connessioni normalmente bloccate fra aree corticali multisensoriali [discusso in Hubbard 2007]. Il **modello della disinibizione** prevede, infatti, che l'assunzione di droghe o farmaci con effetti psicotropi dovrebbe poter generare una simile disinibizione, ossia produrre appunto effetti sinestesici in chi assume la sostanza. Il problema è che questi effetti non sono veramente confrontabili con quanto sembra succedere nei sinesteti veri e propri, le cui esperienze sembrano spontanee, sempre simili per tutta la vita, e non richiedono l'assunzione di nulla in particolare – anche se richiedono, evidentemente, la presentazione dello stimolo che induce l'esperienza sinestetica.

3. PERCEZIONE MULTISENSORIALE O SUGGESTIONE?

Ma si potrebbe chiedere, siamo proprio sicuri che questi resoconti introspettivi corrispondano a esperienze percettive vere e proprie? Quando una persona dice di vedere lampi di luce colorata ogni volta che sente certi suoni, li vede davvero o in qualche modo li immagina? Insomma, la sinestesia è proprio percezione o piuttosto una specie di autosuggestione? In fondo, gli unici dati a favore dell'esistenza del fenomeno sono i resoconti introspettivi dei sinesteti, autodichiaratisi tali. Per quale motivo dovremmo credere loro? Ci sono molti buoni motivi. Innanzi tutto, anche se le associazioni sinestetiche sono largamente idiosincratiche nei loro aspetti qualitativi (ad esempio, nell'associare a un tono un determinato colore piuttosto che un altro), alcune caratteristiche della sinestesia sembrano essere generali. Tutti i sinesteti riferiscono che le loro esperienze sembrano essere immediate, senza sforzo apparente, proprio come avviene nella percezione «normale». Una seconda caratteristica comune è che le esperienze sinestetiche sembrano localizzarsi fenomenicamente nello spazio esterno vicino al corpo, il che potrebbe forse indicare una relazione con meccanismi coinvolti nella percezione dello spazio peripersonale (di cui parleremo nel cap. 9). Inoltre le associazioni sinestetiche sono, pur nella loro idiosincrasia, molto stabili nel tempo. Ad esempio, una ricerca ha esaminato le associazioni tono-colore in un gruppo di sinesteti una prima volta e poi a distanza di un anno [Baron-Cohen *et al.* 1993], trovando che quasi il 95% delle associazioni erano identiche. Infine, le qualità fenomeniche esperite nella sinestesia sono sempre caratteristiche semplici: colori, caratteristiche della trama, forme tridimensionali. Non sono mai stati descritti casi di sinestesie in cui vengano evocate percezioni di veri e propri oggetti inseriti in un contesto: cose come una scena visiva, o un panorama. Sembra naturale quindi concludere che le associazioni sinestetiche coinvolgano le caratteristiche percettive di base che vengono fuse nella percezione multisensoriale di oggetti, e che in qualche modo riguardino una fase del processo percettivo che precede la fusione multisensoriale.

Una seconda serie di prove a favore della natura genuinamente percettiva della sinestesia viene inoltre da una serie di esperimenti molto ingegnosi del famoso neuropsicologo Vilayanur Ramachandran [in parte ispirati da un precedente lavoro di Mills, Boteler e Oliver 1999]. Considerate la figura 7.1. A sinistra c'è una matrice di lettere nere. Se siete come gli autori di questo libro (che non sono sinesteti), tenderete spontaneamente a vedere che le lettere formano delle righe. Si tratta del principio di organizzazione per prossimità, di cui abbiamo già parlato nel capitolo 4. Ma il discorso cambia se avete una **sinestesia grafema-co-**

3 8 3 8   3 8 3 8
7 0 7 0   7 0 7 0
3 8 3 8   3 8 3 8
7 0 7 0   7 0 7 0

F F P F F P P H P P F P H P H F H H H H F P P F P

fig. 7.1. Esperimenti molto eleganti testimoniano degli effetti genuinamente percettivi della sinestesia. Ad esempio, se si presenta una matrice di numeri come quella a sinistra a una persona senza sinestesia, tenderà a percepirla organizzata in righe orizzontali (principio di prossimità). Per un sinesteta numero-colore, prevale invece l'organizzazione in colonne verticali (principio di somiglianza), proprio come se l'individuo in questione «vedesse» i numeri colorati nella maniera in cui li vediamo tutti nella matrice al centro. Se si presenta la figura di destra a persone senza sinestesia, queste hanno notevole difficoltà a percepire un triangolo formato da lettere «H». Per un sinesteta grafema-colore, la percezione del triangolo sembra invece essere immediata, proprio come se le H emergessero grazie a un colore diverso.

fonte: Ridisegnata da Ramachandran e Hubbard [2001].

lore, ossia, se ogni volta che vedete la cifra 3 o 7 la vedete verde, e ogni volta che vedete la cifra 8 o 0 la vedete rossa. Per voi la matrice dovrebbe apparire come nella versione colorata al centro, dove l'organizzazione spontanea non dipende più dalla sola prossimità ma anche dal principio di somiglianza, generando una tendenza a vedere colonne e non più righe. Ramachandran e Hubbard [2001] hanno studiato due sinesteti grafema-colore e, costruendo matrici di lettere appropriate per le loro idiosincratiche associazioni, sono stati in grado di dimostrare che l'organizzazione percettiva spontanea dei due sinesteti era effettivamente modificata come ci si aspetterebbe se, oltre alla prossimità, vi fosse un'influenza della somiglianza fra colori. E non solo. Sfruttando l'associazione fra altri tipi di grafema (lettere) e i colori, Ramachandran e Hubbard hanno anche dimostrato che uno dei due sinesteti era in grado di rilevare immediatamente una forma mascherata da un campo di distrattori, come il triangolo formato dalle H nella parte più a destra della figura 7.1. Per un individuo senza sinestesia, le lettere H, F e P sono invece percettivamente molto simili, e generalmente difficili da distinguere in una disposizione come quella della figura. Rilevare il triangolo risulta dunque molto difficile, come potete verificare voi stessi cercandolo nella figura (sempre se non siete sinesteti).

Infine, in un lavoro più recente Ramachandran e Azoulai [2006] hanno dimostrato che i colori associati alle lettere possono produrre addirittura la per-

cezione di movimento apparente. In quest'ultimo studio, questi autori hanno presentato in sequenza due immagini costruite con le cifre 2 e 5. Le posizioni dei 2 erano scelte a caso in entrambe le immagini, mentre i 5 erano sistemati lungo una linea verticale. Nella prima immagine questa linea era nella zona sinistra mentre nella seconda immagine veniva spostata nella zona destra. Presentando in sequenza le due immagini, i partecipanti senza sinestesia riportavano di vedere solo movimenti casuali in tutte le direzioni. Infatti, le cifre 2 e 5 sono difficili da distinguere (un po' come le F, le P e le H della fig. 7.1) e questo maschera la presenza dei 5 allineati. Non così per un sinesteta grafema-colore, che vedeva i 5 rossi e i 2 verdi. Per questa persona, la linea (rossa) di 5 si segregava spontaneamente dallo sfondo di 2 (verdi) e la successione delle due immagini produceva una vivida impressione di movimento apparente della barra (rossa) verso destra.

Risultati come questi sono assai difficili da spiegare se non ammettendo che per i sinesteti, l'associazione del colore al grafema produce un **effetto genuinamente percettivo**. Se si considera la loro prestazione, è proprio come se vedessero numeri colorati. Va notato però che questi dati si riferiscono a una situazione che non è propriamente multisensoriale: si tratta di associazioni tutte interne al canale visivo, anche se riguardano meccanismi che sappiamo distinti per la codifica del colore e della forma. Allo stato attuale, possiamo quindi essere un po' meno sicuri della natura genuinamente percettiva delle sinestesie che coinvolgono canali diversi, come la sinestesia tono-colore o il caso descritto da Cytowic di sinestesia gusto-forma. Ma la maggior parte dei ricercatori tende oggi a ritenere che anche in questi casi ci siano pochi dubbi in proposito. Inoltre, ci sono dati sperimentali recentissimi in cui la sinestesia sembra avere un effetto su processi indubbiamente multisensoriali. Vediamo di che si tratta.

Il primo di questi due lavori [Parise e Spence 2009] ha studiato una forma di **sinestesia debole**, di cui possiamo avere esperienza tutti anche se non siamo sinesteti veri e propri. L'idea di base è una vecchissima osservazione del grande psicologo gestaltista Wolfgang Köhler [1929] e diventata famosa col nome di fenomeno *takete-maluma*. Osservate la figura 7.2, una riproduzione delle figure originali apparse sul libro di Köhler. Köhler notava che se si chiede di accoppiare le configurazioni con le parole *takete* e *maluma*, nessuno esita: *maluma* è la configurazione curvilinea, *takete* quella spigolosa. È come se il suono dolce e continuo della M e della L si raggruppasse spontaneamente con il contorno senza bruschi cambi di direzione della configurazione curva, e viceversa per il suono a scatti e la configurazione a spigoli. Una specie di organizzazione per somiglianza multisensoriale, insomma, anche se non è del tutto chiaro quali

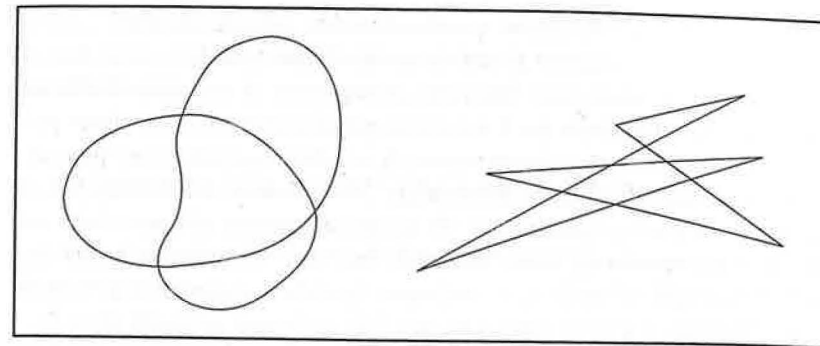



fig. 7.2. Due configurazioni, *takete* e *maluma*. Qual è *takete*? Qual è *maluma*?

fonte: Köhler [1929].

proprietà fisiche dello stimolo sonoro e visivo possano essere considerate simili e collegabili alle proprietà percepite. In questa forma di sinestesia debole [Martino e Marks 2001], dunque, si ha esperienza di un'associazione sinestetica, spontanea e non appresa, fra stimoli diversi.

Naturalmente si tratta di una cosa diversa dalla sinestesia vera e propria, in cui uno stimolo evoca la percezione di qualcosa d'altro per cui uno stimolo non c'era. Ma si tratta comunque di un fenomeno rilevante, soprattutto se ipotizziamo che la sinestesia possa essere considerata un continuum di fenomeni associativi, da casi più deboli come questi ai casi estremi in cui l'associazione addirittura genera un percetto nuovo. Per la loro ricerca, Parise e Spence hanno utilizzato una serie di stimoli visivi e acustici che possono associarsi spontaneamente, come nel fenomeno *takete-maluma*. Ad esempio, un'associazione altrettanto spontanea è quella fra configurazioni relativamente piccole e suoni ad alta frequenza e configurazioni relativamente grandi e suoni a bassa frequenza. Sfruttando questa osservazione, Parise e Spence hanno accoppiato suoni alti o bassi con figure grandi o piccole, generando coppie che tendono ad associarsi spontaneamente e coppie che non hanno una simile tendenza. Armati di queste coppie di stimoli, le hanno presentate lievemente sfasate nel tempo o in diverse posizioni spaziali, chiedendo ai partecipanti, che non erano sinesteti veri e propri ma normalissimi studenti Universitari, di dare un giudizio di ordine temporale o di posizione relativa.

I risultati hanno dimostrato che la precisione dei giudizi era peggiore nel caso delle coppie che tendevano spontaneamente ad associarsi rispetto a quelle che non avevano questa caratteristica. Detto in altra maniera, i partecipanti

erano più incerti nei loro giudizi quando dovevano valutare una coppia suono basso-figura grande rispetto ai casi suono alto-figura grande o suono basso-figura piccola. Se accettiamo l'idea che questa forma di sinestesia debole sia comunque rilevante anche per il fenomeno più drammatico da cui siamo partiti, abbiamo dunque una dimostrazione di un effetto genuinamente percettivo anche in una situazione multisensoriale. Ma l'interesse del risultato non si esaurisce con questa constatazione. Se ripensate a quanto abbiamo detto nel capitolo 4 a proposito dei limiti del modello bayesiano, ricorderete che uno dei problemi cruciali del modello sta nel capire quando due segnali unisensoriali vanno integrati, e quando no, e che una delle soluzioni proposte introduce l'idea di un *interaction prior* che determina appunto la probabilità che i due segnali vengano da una fonte comune [Kording *et al.* 2007]. I risultati di Parise e Spence hanno il merito di dimostrare che tale probabilità potrebbe dipendere da associazioni simili a quelle del fenomeno *takete-maluma*, ossia da proprietà relazionali della struttura dello stimolo, e non solo da associazioni arbitrarie apprese.

Il secondo risultato, anch'esso recentissimo, è dovuto di nuovo al gruppo di Gemma Calvert a Oxford [Bargary *et al.* 2009]. Questi ricercatori hanno studiato 9 individui con **sinestesia parola-colore** utilizzando stimoli linguistici audiovisivi. Come forse ricorderete, perché ne abbiamo già accennato nell'introduzione, quando lo stimolo visivo (movimento delle labbra del parlante) e quello acustico (suono) non coincidono, in molti casi il suono linguistico percepito è una specie di compromesso fra i segnali unisensoriali discordanti. Si tratta del notissimo effetto McGurk, di cui ci occuperemo ancora più avanti. Ad esempio, una variante dell'effetto McGurk è questa: se si vede una persona articolare la parola *peat* («torba») mentre si sente il suono *neat* («pulito»), viene percepita la parola *meat* («carne»). I ricercatori di Oxford hanno trovato alcuni sinesteti per cui parole come queste evocavano sistematicamente alcuni colori. Per uno di questi, ad esempio, *peat* evoca l'azzurro, *neat* evoca il verde chiaro e *meat* il blu scuro. I ricercatori hanno quindi provato a presentare la parola *neat* nel solo canale acustico, oppure accoppiata a *peat* in quello visivo. Ebbene: quando veniva presentata la sola *neat* acustica, il sinesteta riportava di vedere il verde; ma quando invece *neat* era accoppiato a *peat* visiva, il sinesteta riportava il blu scuro. La stessa cosa accadeva per gli altri sinesteti, naturalmente con altri colori in funzione delle loro idiosincratiche associazioni. Dunque, nella condizione di conflitto audiovisivo, le associazioni parola-colore coinvolgevano i colori associati alla parola audiovisiva «illusoria», non quelli associati alle parole effettivamente presenti nei segnali unisensoriali. Questo suggerisce che il meccanismo sinestetico agisce a un livello successivo all'integrazione dei segnali unisensoriali, documentando un effetto della sinestesia su processi sicuramente multisensoriali.

4. SINESTESIA E BASI NEURALI DELLA MULTISENSORIALITÀ

Molti ricercatori ritengono che la sinestesia non sia solo una curiosità, ma possa fornirci indicazioni preziose su come funziona la percezione multisensoriale e sulle sue basi neurali. Questo non solo per le caratteristiche del fenomeno ma anche per la sua diffusione. Se, infatti, la sinestesia era considerata inizialmente una condizione piuttosto rara, riguardante non più di una persona su 25.000 [Cytowic e Wood 1982b], secondo stime più recenti e statisticamente affidabili circa 1 persona su 23 avrebbe una qualche forma di sinestesia [Simner *et al.* 2006]. Inoltre, mentre i primi tentativi di spiegazione mettevano l'accento sul possibile ruolo del sistema limbico, oggi abbiamo delle prove piuttosto convincenti che la base neurale delle esperienze sinestesiche coinvolge aree della corteccia. Questo lo sappiamo grazie a studi di neuroimmagine, anche se è stato notato che l'interpretazione di questi è spesso difficile per il numero insufficiente di sinesteti studiati nei singoli esperimenti [Hubbard 2007], e grazie a ricerche con particolari tecniche psicofisiche [Palmeri *et al.* 2002] che hanno dimostrato, in sinesteti grafema-colore, che le luci colorate sinestetiche potevano essere elicitate anche da grafemi rilevabili dal sistema percettivo solo dopo la fusione binoculare, che avviene appunto a livello corticale.

Anche se una spiegazione definitiva è ancora lontana, un'ipotesi che gode di un certo credito si basa sull'idea che il fenomeno dipenda dal persistere in età adulta di connessioni cortico-corticali che vengono di norma disattivate durante lo sviluppo. Secondo l'**ipotesi della sinestesia neonatale** [Maurer 1993], ad esempio, fino a circa quattro mesi di età i neonati avrebbero esperienze sensoriali ampiamente indifferenziate. Nel corso dello sviluppo successivo, i canali sensoriali diventerebbero invece più modulari, producendo una sempre maggiore differenziazione delle esperienze percettive associate alla loro stimolazione. Nei sinesteti, questo processo di differenziazione avverrebbe in modo incompleto, preservando l'efficacia di alcune connessioni [Baron-Cohen *et al.* 1993]. Questo processo potrebbe riguardare tutti gli individui, o forse solo una sottopopolazione geneticamente predisposta. Indagini sulla prevalenza familiare della sinestesia infatti suggeriscono una possibile componente genetica della stessa [Baron-Cohen *et al.* 1996; Simner *et al.* 2006].

Nel definire la tipologia delle possibili associazioni sinestetiche potrebbe inoltre avere un ruolo la **topologia delle aree corticali**. Infatti Ramachandran e Hubbard [2001] hanno fatto notare che l'area occipitale ritenuta responsabile dell'elaborazione del colore (area V4) si colloca nel cervello nella zona immediatamente adiacente all'area responsabile del riconoscimento di lettere e numeri, nel giro fusiforme del lobo temporale. Questa caratteristica potrebbe aumentare la possibilità di avere connessioni fra queste due aree, le quali, se non diventano completamente modulari nel corso dello sviluppo, potrebbero risultare in una residua connettività anomala, in grado di generare i percetti sinestesici. Il principio potrebbe valere anche per altri tipi di sinestesia più propriamente multisensoriale. Ad esempio, l'area deputata all'elaborazione del gusto è vicina alle aree somatosensoriali primarie, il che potrebbe favorire la possibilità di sinestesie gusto-forma come quella descritta all'inizio del paragrafo. Come fa notare Hubbard [2007], la prossimità delle aree corticali potenzialmente coinvolte non va considerata una condizione necessaria, ma solo un fattore facilitante. L'ipotesi richiede infatti solo che siano presenti delle connessioni, non che queste debbano essere necessariamente fra aree vicine. Attraverso lo studio delle differenze anatomiche fra sinesteti e non sinesteti, lo studio dei fenomeni sinestetici potrebbe dunque contribuire a fornire dati sui circuiti neurali coinvolti nella percezione multisensoriale di oggetti.

PARTE TERZA

Lo spazio